Performance Investigation of Some Existing Numerical Methods for Inverse Problems

by

CHEUNG, Man Wah

A Thesis Submitted in Partial Fulfilment
of the Requirements for the Degree of
Master of Philosophy
in
Mathematics

June 2007

Thesis/Assessment Committee

Professor CHAN Hon Fu Raymond (Chair)
Professor ZOU Jun (Thesis Supervisor)
Professor AU Kwok Keung Thomas (Committee Member)
Professor MU Mo (External Examiner)

Abstract

Inverse problems arise in many areas of science and engineering. An inverse problem is often ill-posed, i.e., its solution does not satisfy one or more of the three criteria: existence, uniqueness, and stability. We are particularly interested in the instability, i.e., the solution does not depend continuously on the observation data. Since most inverse problems cannot be solved analytically, computational methods play an important role in solving inverse problems and this is the main objective of this thesis.

This thesis is divided into two parts. The first part is to review the major properties of inverse problems and the basic theory of regularization methods. We will focus on the Fredholm integral equation of the first kind, which is one of the most classical examples of inverse problems. The second part is to present some of the existing numerical methods for solving inverse problems. Three of the most important regularization methods: Tikhonov regularization, Landweber iteration and truncated singular value decomposition method, and multilevel methods will be studied and analyzed. The performance of these methods will also be investigated through numerical experiments.

摘要

反問題在科學及工程中廣泛出現。反問題通常是不適定的，即問題的解不附合以下三個條件中的一個或多個：存在性、唯一性、及穩定性。穩定性尤其值得我們關注：問題的解不連續依賴於觀察數據。由於大多數的反問題都無法得到解析解，因此計算方法在解反問題中佔有一非常重要的角色，這亦是本論文的主要探討範疇。

這論文主要分爲兩部分。第一部分爲綜述反問題的主要性質及正規化方法的基本理論。我們會集中討論反問題其中一種最經典的例子：第一類 Fredholm 積分方程。第二部分爲闡述解反問題的一些現有數值方法。我們會研究及分析其中三種最重要的正規化方法：Tikhonov 正規化、Landweber 迭代法、及截斷的奇異值分解方法，和多水平方法。最後，我們會透過數值實驗去研究這些不同方法的表現。

ACKNOWLEDGMENTS

I wish to express my sincere gratitude to my supervisor, Prof. Jun Zou, for his guidance and support throughout these two years of my postgraduate study.

I would also like to thank Jianfeng Cai, Bangti Jin, Jingzhi Li, Hongyu Liu, Yuliang Wang, Tim Yin Shung Wong, Yifeng Xu and Kai Zhang for the discussions on theoretical analysis, programming problems and subtleties of LaTeX, and all their help on my work and study.

I would like to thank Prof. H. Brunner, Prof. X.C. Cai, Prof. K. Chen, Prof. P.G. Ciarlet, Prof. S.H. Lui and Prof. X.P. Wang for their teaching on different topics on applied mathematics. Thanks are also given to all the professors who have taught me in my undergraduate study. Special thanks are given to my secondary teachers, especially Mr. Lap Foo Chu and Mr. Siu Wai Fan for their teaching on mathematics.

Finally I wish to thank the people who have loved me and cared about me in my life.

To my father and mother

Contents

List of Figures

List of Tables

Chapter 1

Introduction to Inverse Problems

The research field of inverse problems has experienced an explosive growth in the last few decades. It received a great deal of attention by applied mathematicians, statisticians, engineers and physical scientists due to its importance of applications like medical imaging, geological prospecting and image processing (see [2, 7, 16] for more examples). This is also due to the recent development of powerful computers and fast, reliable numerical methods with which to carry out the numerical simulation process. In this chapter, we will first give a brief introduction to inverse problems, followed by some of their major properties. Then we will present some typical examples, and finally the outline of this thesis will be given.

1.1 Major properties

A direct problem is the one in which we determine the effect y of a given cause x when a definite mathematical model K is available, i.e., we are given x and K, and determine $y = Kx$. There are mainly two kinds of inverse problems. The first one is that we are given y and K, and solve $Kx = y$ for x (causation). The second one is that we are given y and x, and determine K (model identifi-

cation). There is a fundamental difference between direct problems and inverse problems. *Direct problems are usually well-posed* in the sense of Hadamard [8], while *inverse problems are usually ill-posed.* Hadamard introduced the concept of well-posedness in the sense that it possesses the following three properties:

1. *There exists a solution of the problem for any given set of data.*
2. *There is at most one solution of the problem for any given set of data.*
3. *The solution depends continuously on the data, i.e., a small perturbation in the data will lead to a small deviation in the solution.*

The three properties are referred to as the property of *existence, uniqueness* and *stability* respectively. To formulate the notion of well-posedness mathematically, we quote the definition from [16] below.

Definition 1.1.1 *Let X and Y be two normed spaces, $K : X \to Y$ a (not necessarily linear) mapping. The equation $Kx = y$ is called well-posed if the following three criteria hold:*

1. ***Existence:*** *For every $y \in Y$, there is (at least one) $x \in X$ such that $Kx = y$.*
2. ***Uniqueness:*** *For every $y \in Y$, there is at most one $x \in X$ with $Kx = y$.*
3. ***Stability:*** *The solution x depends continuously on y, i.e., for every sequence $\{x_n\} \subset X$ with $Kx_n \to Kx$ $(as\ n \to \infty)$, it follows that $x_n \to x$ $(as\ n \to \infty)$.*

Equations $Kx = y$ for which (at least) one of these properties does not hold are called ill-posed.

Often, existence and uniqueness can be forced by enlarging or reducing the solution space. Therefore, stability is our main concern. In practice, if a problem does not possess the property of stability, then the solution cannot be computed by numerical methods reliably due to errors in measurement and computation.

1.2 Typical examples

In this section, we give some typical examples of inverse problem together with its direct problem.

Example 1.1 (Computerized tomography, see [9])
The direct problem is to determine the damping of the x-rays given the x-ray source and the object being scanned. The inverse problem is to reconstruct the object being scanned from information about the locations of the x-ray sources and measurements of their damping.

Example 1.2 (Inverse scattering problem, see [16])
The direct scattering problem is to determine the scattered field from a knowledge of an incident field and a given scattering object. The inverse problem is to find the shape of a scattering object, given the field scattered by this object.

Example 1.3 (Backward heat equation, see [16])
Consider the 1-D heat equation

$$\frac{\partial u(x,t)}{\partial t} = \frac{\partial^2 u(x,t)}{\partial x^2}$$

with boundary conditions

$$u(0,t) = u(\pi,t) = 0, \qquad t \geq 0,$$

and initial condition

$$u(x,0) = u_0(x), \qquad 0 \leq x \leq \pi.$$

The direct problem is to solve the classical initial boundary value problem: Given the initial temperature distribution u_0 and the final time T, determine $u(\cdot, T)$. For the inverse problem, one measures the final temperature distribution $u(\cdot, T)$ and tries to determine the temperature at earlier time $t < T$, e.g. the initial temperature $u(\cdot, 0)$.

Example 1.4 (Image deblurring, see [26])

Consider the Fredholm first kind integral equation of convolution type

$$\int_0^1 \int_0^1 k(x-x', y-y') f(x', y')\, dx'\, dy' = g(x, y), \qquad 0 < x, y < 1.$$

The function f represents light source intensity and g represents image intensity. The kernel k characterizes blurring effects that occur during image formation. The direct problem is to determine the blurred image g given the source f and the kernel k. The inverse problem is to determine the source f (given the kernel k) from the observation of the image g.

Example 1.5 (Exponential growth model, see [7])

Consider the differential equation

$$\frac{du}{dt} = ru.$$

Here $u(t)$ might represent the population of a colony of bacteria and $r(t)$ represents the growth rate. The direct problem is to solve the differential equation for some given function $r(t)$. For the inverse problem, one observes the quantity u and tries to determine the growth rate r.

Example 1.6 (Diffusion in inhomogeneous medium, see [16])

The equation of diffusion in an inhomogeneous medium is described by the equation

$$\frac{\partial u(x,t)}{\partial t} = \frac{1}{c} \nabla \cdot (\kappa \nabla u(x,t)), \qquad x \in D,\ t > 0,$$

where c is a constant and $\kappa = \kappa(x)$ is a parameter describing the medium. The direct problem is to solve the boundary value problem for this equation given the boundary conditions $u|_{\partial D}$ and the function $\kappa(x)$. For the inverse problem, one measures u and the flux $\frac{\partial u}{\partial \nu}$ on the boundary ∂D and tries to determine the unknown function κ in D.

Examples 1.5 and 1.6 are examples of *parameter identification problems* in differential equations.

1.3 Thesis outline

The rest of this thesis is organized as follows. Chapter 2 is a review of some operator theory including spectral theory and compact operator theory. The general regularization theory as well as some concrete regularization methods for solving Fredholm integral equations of the first kind will be given in chapter 3. In chapter 4, the multilevel algorithms proposed by King in [14] will be presented. Finally, some numerical results of the regularization methods in chapter 3 and the multilevel methods in chapter 4 will be given in chapter 5.

Chapter 2

Some Operator Theory

In this chapter, we will first introduce one of the most classical examples of inverse problems - Fredholm integral equations of the first kind. Next we will review some results in operator theory including spectral theory and compact operator theory. Finally we will introduce the singular system and the Moore-Penrose generalized inverse of a compact linear operator, which are very useful in the analysis of the "generalized" solutions of the first kind Fredholm integral equations.

2.1 Fredholm integral equation of the first kind

A *Fredholm integral equation of the first kind* is an equation of the form

$$\int_a^b k(s,t)u(t)\,dt = g(s) \tag{2.1}$$

where g and $k(\cdot,\cdot)$ are given functions and u is the solution that we want to seek. The function g is usually called the "data" and $k(\cdot,\cdot)$ is called the *kernel* of the equation. When the variable t represents time and the past is unaffected by the future, then $k(s,t) = 0$ for $s < t$, and (2.1) becomes

$$\int_a^s k(s,t)u(t)\,dt = g(s), \tag{2.2}$$

which is the form of a *Volterra integral equation of the first kind.* Fredholm integral equations of the first kind is usually an ill-posed problem [6]. First, let us consider the existence of solutions. For example, if the kernel k is continuous and u is integrable, then the function g defined by (2.1) is also continuous. So if the given function g is not continuous while the kernel is, then (2.1) has no integrable solution. An other point is the uniqueness of solutions. For example, let us consider the equation

$$\int_0^\pi k(s,t)u(t)\,dt = 0$$

where $k(s,t) = s\sin t$. Of course $u(t) \equiv 0$ is a solution, but so is each of the following functions

$$u_n(t) = \sin nt, \qquad n = 2, 3, \ldots$$

due to the orthogonality relations. The last point that we want to mention here is the stability of the solutions. From the *Riemann-Lebesgue lemma*, we know that if $k(\cdot,\cdot)$ is any square integrable function, then

$$\int_0^\pi k(s,t)\sin nt\,dt \to 0 \quad \text{as} \quad n \to \infty.$$

It follows that if $u(t)$ is a solution of

$$\int_0^\pi k(s,t)u(t)\,dt = g(s),$$

then

$$\int_0^\pi k(s,t)(u(t) + M\sin nt)\,dt \to g(s) \quad \text{as} \quad n \to \infty$$

for any $M > 0$. Therefore if we take M very large, then for very large n, a small perturbation in the data

$$\tilde{g}(s) = g(s) + \int_0^\pi k(s,t)M\sin nt\,dt$$

is accounted for by a large change in the solution

$$\tilde{u}(t) = u(t) + M\sin nt.$$

Generally, the solutions depend discontinuously on the data (please refer to [7] for more examples).

Consider the Fredholm integral equation of the first kind in an abstract form:

$$Ku = g \tag{2.3}$$

where K is a linear operator from some topological space X to a topological space Y. Then the conditions of well-posedness can be said in another way. Existence means $Y = K(X)$, or the mapping K is surjective. Uniqueness is simply the injectivity of K. If the conditions of existence and uniqueness are both satisfied, the inverse operator $K^{-1} : Y \to X$ exists and the condition of stability is equivalent to the continuity of K^{-1}.

2.2 Compact operator theory

In this section, we will first establish some notation, then review some results in operator theory including *spectral theory* and *compact operator theory*, which can be found in most books on functional analysis (e.g. [5, 18]). Here we use $\langle \cdot, \cdot \rangle$ to denote the inner product in a Hilbert space and $\| \cdot \|$ to denote the norm generated by the inner product. If S is a subset of a Hilbert space H, then $S^{\perp}$ denotes the *orthogonal complement* of S, i.e.

$$S^{\perp} = \{y \in H : \langle x, y \rangle = 0 \ \ \forall x \in S\}.$$

If T is a continuous linear operator from a Hilbert space H_1 to a Hilbert space H_2, then $T^* : H_2 \to H_1$ denotes the *adjoint* of T and is defined by

$$\langle Tx, y \rangle = \langle x, T^* y \rangle \qquad \text{for all } \ x \in H_1, y \in H_2.$$

The *range* $R(T)$ and the *null space* $N(T)$ of T are defined by

$$R(T) = \{Tx : x \in \mathcal{D}(T)\}$$

and

$$N(T) = \{x \in \mathcal{D}(T) : Tx = 0\}$$

respectively where $\mathcal{D}(T)$ denotes the *domain* of T. The following theorem [5] relates the range and null space of T to that of T^*.

Theorem 2.2.1 *If $T : H_1 \to H_2$ is a continuous linear operator, then*

$$R(T)^\perp = N(T^*) \quad \textit{and} \quad N(T)^\perp = \overline{R(T^*)}.$$

Proof For the first equality, if $y \in N(T^*)$, then for any $x \in H_1$,

$$\langle Tx, y\rangle = \langle x, T^*y\rangle = 0$$

and hence $y \in R(T)^\perp$. Moreover, if $y \in R(T)^\perp$, the same equation shows that $y \in N(T^*)$. Hence $R(T)^\perp = N(T^*)$.

On the other hand, replacing T by T^* in the first equality and using the fact that $T^{**} = T$, we have $R(T^*)^\perp = N(T^{**}) = N(T)$, and so

$$N(T)^\perp = R(T^*)^{\perp\perp}.$$

Since $R(T^*)^{\perp\perp} = \overline{R(T^*)}$, the second equality follows. □

Replacing T by T^*, we also have

$$R(T^*)^\perp = N(T) \quad \text{and} \quad N(T^*)^\perp = \overline{R(T)}.$$

Note that for linear operators continuity is equivalent to boundedness, i.e. the finiteness of the number

$$\|T\| = \sup\{\|Tx\| : \|x\| = 1\},$$

called the *norm* of T. Moreover, if T is bounded, then

$$\|T\| = \|T^*\| = \|TT^*\|^{1/2}.$$

The *spectrum* of a linear operator $T : H \to H$ is the set of complex numbers $\sigma(T)$ defined by

$$\sigma(T) = \{\lambda \in \mathbf{C} : T - \lambda I \text{ has no bounded inverse}\}$$

where I is the identity operator on H. The *spectral radius* of T is the real number $|\sigma(T)|$ defined by

$$|\sigma(T)| = \sup\{|\lambda| : \lambda \in \sigma(T)\}.$$

An operator $T : H \to H$ is called *self-adjoint* if $T = T^*$. If T is a bounded self-adjoint linear operator, then

$$\|T\| = \sup\{|\langle Tx, x\rangle| : \|x\| = 1\}.$$

In this case, the spectral bounds of T are defined by

$$m_T = \inf\{\langle Tx, x\rangle : \|x\| = 1\} \quad \text{and} \quad M_T = \sup\{\langle Tx, x\rangle : \|x\| = 1\}$$

respectively. The spectrum $\sigma(T)$ is a nonempty subset of the interval $[m_T, M_T]$ and $m_T, M_T \in \sigma(T)$. We have the following theorem [5] relating the spectral radius to the norm of T.

Theorem 2.2.2 *If $T \in B(H, H)$, then $|\sigma(T)| \leq \|T\|$. Moreover, if T is self-adjoint, then $|\sigma(T)| = \|T\|$.*

Proof If $|\lambda| > \|T\|$, then $\|\lambda^{-1}T\| < 1$ and therefore $(I - \lambda^{-1}T)^{-1} \in B(H, H)$. But then $(\lambda I - T)^{-1} = \lambda^{-1}(I - \lambda^{-1}T)^{-1} \in B(H, H)$ which implies that λ is not in $\sigma(T)$. So $|\lambda| \leq \|T\|$ for all $\lambda \in \sigma(T)$ and hence $|\sigma(T)| \leq \|T\|$.

If T is self-adjoint, then $m_T, M_T \in \sigma(T)$. Therefore $|\sigma(T)| \geq \max\{|m_T|, |M_T|\} = \|T\|$ and hence $|\sigma(T)| = \|T\|$. □

A complex number λ is called an *eigenvalue* of T if there is some nonzero vector x, called the *eigenvector associated with* λ, such that $Tx = \lambda x$. Every eigenvalue of T is a member of $\sigma(T)$. If T is self-adjoint, then the eigenvectors associated

with distinct eigenvalues are orthogonal since, if $Tx = \lambda x$ and $Ty = \mu y$ with $\lambda \neq \mu$, then

$$\lambda\langle x, y\rangle = \langle Tx, y\rangle = \langle x, Ty\rangle = \mu\langle x, y\rangle$$

which implies $\langle x, y\rangle = 0$. A bounded linear operator K from a normed linear space X into a normed linear space Y is called *compact* if for each bounded set B in X, the set $K(B)$ has compact closure in Y. From the definition, it is clear that the composition of a compact operator with a bounded operator is also compact. Consider the operator K defined as

$$Ku(s) = \int_a^b k(s,t)u(t)\,dt.$$

If $k(\cdot,\cdot)$ is square integrable over $[c,d] \times [a,b]$, then K is a compact operator from $L^2[a,b]$ to $L^2[c,d]$. Compact self-adjoint operators have a very nice spectrum [6]:

1. *each nonzero member of the spectrum is an isolated point which is an eigenvalue of the operator;*

2. *for each nonzero eigenvalue λ, the eigenspace associated with λ, i.e. the set $N(K - \lambda I)$, is finite-dimensional;*

3. *the eigenvalues form a sequence $\lambda_1, \lambda_2, \ldots$, which (if infinite) converges to zero.*

If we repeat each eigenvalue in this list according to the dimension of its associated eigenspace, we may form a sequence $x_1, x_2, \ldots$ of associated orthonormal eigenvectors. Here we state the *spectral theorem* [6] for a compact self-adjoint operator.

Theorem 2.2.3 *If $K : H \to H$ is a compact self-adjoint linear operator with eigenvalues $\lambda_1, \lambda_2, \ldots$ (repeated according to the dimension of the associated eigenspace) and associated orthonormal eigenvectors $x_1, x_2, \ldots$, then for any $u \in H$,*

$$Ku = \sum_n \lambda_n \langle u, x_n\rangle x_n.$$

With the above theorem, we can define functions of a compact self-adjoint operator K in the following way. Given a real-valued continuous function f on $\sigma(K)$, we define $f(K)$ by

$$f(K)u = \sum_n f(\lambda_n)\langle u, x_n\rangle x_n,$$

then we have $\sigma(f(K)) = f(\sigma(K))$. The operator $f(K)$ so defined is self-adjoint and compact, and by using Theorem 2.2.2,

$$\|f(K)\| = |\sigma(f(K))| = \sup\{|f(\lambda)| : \lambda \in \sigma(K)\}.$$

2.3 Singular system

Suppose $K : H_1 \to H_2$ is a compact linear operator. Then $K^*K : H_1 \to H_1$ is a compact self-adjoint linear operator and any eigenvalue λ of K^*K satisfies

$$\lambda\langle x, x\rangle = \langle K^*Kx, x\rangle = \langle Kx, Kx\rangle = \|Kx\|^2 \geq 0$$

where x is an eigenvector associated with λ. So we can arrange the non-negative eigenvalues of K^*K as $\lambda_1 \geq \lambda_2 \geq \cdots$ and let $v_1, v_2, \ldots$ be the associated sequence of orthonormal eigenvectors. Then $\{v_j\}$ is a complete orthonormal set for $\overline{R(K^*)} = N(K)^\perp$. Let $\mu_j = \sqrt{\lambda_j}$ and $u_j = \mu_j^{-1}Kv_j$, then

$$K^*u_j = \mu_j v_j \quad \text{and} \quad Kv_j = \mu_j u_j.$$

Moreover,

$$KK^*u_j = \mu_j Kv_j = \mu_j^2 u_j = \lambda_j u_j$$

and $\{u_j\}$ forms a complete orthonormal set for $\overline{R(K)} = N(K^*)^\perp$. The system $\{u_j, v_j; \mu_j\}$ is called a *singular system* for the operator K and the numbers μ_j are called *singular values* of K [7]. Any $f \in H_1$ can be represented as

$$f = Pf + \sum_{j=1}^{\infty}\langle f, v_j\rangle v_j$$

where P is the orthogonal projection operator of H_1 onto $N(K)$ and hence

$$Kf = \sum_{j=1}^{\infty} \mu_j \langle f, v_j \rangle u_j.$$

This representation of the operator K is called the *singular value decomposition (SVD)*. The following theorem [7], which is known as *Picard criterion*, is a result on the existence of solutions of first kind Fredholm integral equations.

Theorem 2.3.1 *Let $K : H_1 \to H_2$ be a compact linear operator with singular system $\{u_j, v_j; \mu_j\}$. The equation $Kf = g$ has a solution if and only if $g \in \overline{R(K)}$ and*

$$\sum_{j=1}^{\infty} \mu_j^{-2} |\langle g, u_j \rangle|^2 < \infty. \tag{2.4}$$

Proof If $Kf = g$ has a solution f, then $g \in R(K) \subset \overline{R(K)}$ and

$$\mu_j^{-2} |\langle g, u_j \rangle|^2 = \mu_j^{-2} |\langle Kf, \mu_j^{-1} K v_j \rangle|^2 = \mu_j^{-2} |\langle f, \mu_j^{-1} \lambda_j v_j \rangle|^2 = |\langle f, v_j \rangle|^2,$$

and hence by Bessel's inequality,

$$\sum_{j=1}^{\infty} \mu_j^{-2} |\langle g, u_j \rangle|^2 = \sum_{j=1}^{\infty} |\langle f, v_j \rangle|^2 \leq \|f\|^2 < \infty.$$

Conversely, if $g \in \overline{R(K)} = N(K^*)^{\perp}$ and if (2.4) holds, then any function of the form

$$f = \sum_{j=1}^{\infty} \mu_j^{-1} \langle g, u_j \rangle v_j + \varphi$$

where $\varphi \in N(K)$, is a solution of $Kf = g$ since g can be represented as

$$g = \sum_{j=1}^{\infty} \langle g, u_j \rangle u_j.$$

□

The condition $g \in R(K)$ may be viewed as an abstract smoothness or regularity condition in the sense that g inherits some of the smoothness of the kernel [6].

From Picard criterion, it can be seen that $|\langle g, u_j \rangle|$, the magnitude of the Fourier coefficient of g with respect to the singular functions u_j, needs to decay fast enough relative to the singular values μ_j (note that $\mu_j \to 0$ as $j \to \infty$) in order to ensure the regularity.

2.4 Moore-Penrose generalized inverse

Consider again equation (2.3). It is well-posed if the inverse operator $K^{-1} : Y \to X$ exists and is continuous. However, the inverse operator K^{-1} does not exist in general. In this section, we introduce a more general concept of solution u, called a *least squares solution*, which is defined as

$$\|Ku - g\| = \inf\{\|Kx - g\| : x \in H_1\}.$$

Then we have the following theorem [4] which shows some equivalent characterizations of least squares solutions.

Theorem 2.4.1 *Suppose $K : H_1 \to H_2$ is a bounded linear operator from a Hilbert space H_1 into a Hilbert space H_2. The following conditions are equivalent:*

(i) $\|Ku - g\| = \inf\{\|Kx - g\| : x \in H_1\}$,

(ii) $K^*Ku = K^*g$,

(iii) $Ku = Pg$,

where P is the orthogonal projection operator of H_2 onto $\overline{R(K)}$.

Proof $(i) \Rightarrow (ii)$: Suppose $\|Ku - g\| = \inf\{\|Kx - g\| : x \in H_1\}$. By using Pythagorean theorem and the fact that $Pg - g \in R(K)^{\perp}$,

$$\begin{aligned} \|Ku - g\|^2 &= \|Ku - Pg\|^2 + \|Pg - g\|^2 \\ &\geq \|Ku - Pg\|^2 + \|Ku - g\|. \end{aligned}$$

The last inequality is due to our assumption. So we have $Ku - g = Pg - g \in R(K)^{\perp} = N(K^*)$ and hence $K^*Ku = K^*g$.

$(ii) \Rightarrow (iii)$: If $K^*Ku = K^*g$, then $Ku - g \in N(K^*) = R(K)^{\perp}$ and

$$0 = P(Ku - g) = Ku - Pg.$$

$(iii) \Rightarrow (i)$: Suppose $Ku = Pg$. Then for any $x \in H_1$, again by using Pythagorean theorem and the fact that $Pg - g \in R(K)^{\perp}$,

$$\begin{aligned} \|Kx - g\|^2 &= \|Kx - Pg\|^2 + \|Pg - g\|^2 \\ &= \|Kx - Pg\|^2 + \|Ku - g\|^2 \\ &\geq \|Ku - g\|^2 \end{aligned}$$

and (i) follows. □

From (iii) we see that equation (2.3) has a least squares solution if and only if $Pg \in R(K)$, i.e. if and only if g is a member of the dense subspace $R(K) + R(K)^{\perp}$ of H_2. Using the above theorem, the set of least squares solution can be written as

$$\{u \in H_1 : K^*Ku = K^*g\}$$

which, by the continuity and linearity of K and K^*, is a closed convex set, and hence this set of least squares solutions has a unique element of minimal norm which is denoted by $K^{\dagger}g$. The operator $K^{\dagger}$ defined on the dense subspace $\mathcal{D}(K^{\dagger}) = R(K) + R(K)^{\perp}$ in this way is called the *Moore-Penrose generalized inverse* of K, and $u^{\dagger} := K^{\dagger}g$ is called the *best approximate solution* of $Ku = g$. The following theorem [4] relates the range and null space of $K^{\dagger}$ to that of K.

Theorem 2.4.2 *Let $K^{\dagger}$ be the Moore-Penrose generalized inverse of K, then*

$$N(K^{\dagger}) = R(K)^{\perp} \quad \textit{and} \quad R(K^{\dagger}) = N(K)^{\perp}.$$

Proof Since

$$g \in N(K^\dagger) \Leftrightarrow K^\dagger g = 0 \Leftrightarrow Pg = K \cdot 0 = 0 \Leftrightarrow g \in R(K)^\perp,$$

the first equality follows. For the second equality, we let $g \in \mathcal{D}(K^\dagger)$. Suppose $K^\dagger g = u_1 + u_2 \in N(K)^\perp \oplus N(K)$. Then u_1 is a least squares solution of $Ku = g$ since

$$Ku_1 = K(u_1 + u_2) = KK^\dagger g = Pg.$$

If $u_2 \neq 0$, by using Pythagorean theorem, we have

$$\|u_1\|^2 < \|u_1\|^2 + \|u_2\|^2 = \|u_1 + u_2\|^2 = \|K^\dagger g\|^2$$

which contradicts that $K^\dagger g$ is the least squares solution of minimal norm. Therefore $K^\dagger g = u_1 \in N(K)^\perp$. Conversely, suppose that $u \in N(K)^\perp$. Let $g = Ku$, then $Ku = PKu = Pg$ and so u is a least squares solution of $Ku = g$. If x is another least squares solution, then $Kx = Pg = Ku$ and so

$$x - u = \overline{u} \in N(K).$$

It follows that by using Pythagorean theorem,

$$\|x\|^2 = \|u\|^2 + \|\overline{u}\|^2 \geq \|u\|^2$$

which implies that u is the least squares solution of minimal norm, i.e. $u = K^\dagger g \in R(K^\dagger)$. Hence $R(K^\dagger) = N(K)^\perp$. □

With the Moore-Penrose generalized inverse $K^\dagger$, the questions of existence and uniqueness are settled. For stability, recall that for linear operators, continuity is equivalent to boundedness. It can be shown that $K^\dagger$ is bounded if and only if $R(K)$ is closed [2]. It turns out that for compact operator K, the Moore-Penrose generalized inverse $K^\dagger$ is bounded if and only if $R(K)$ is finite dimensional [6]. Therefore the first kind equation (2.1) is well-posed only if the kernel k is degenerate, i.e. has the form

$$k(s,t) = \sum_{i=1}^{n} \varphi_i(s)\psi_i(t).$$

It is therefore necessary to devise some means of imposing stability when solving first kind Fredholm integral equations. At last we state the following theorem [6] which gives an explict representation of the Moore-Penrose generalized inverse of a compact operator to end this section.

Theorem 2.4.3 *If $K : H_1 \to H_2$ is a compact linear operator with singular system $\{u_j, v_j; \mu_j\}$ and $g \in \mathcal{D}(K^\dagger)$, then*

$$K^\dagger g = \sum_{j=1}^{\infty} \frac{\langle Pg, u_j \rangle}{\mu_j} v_j = \sum_{j=1}^{\infty} \frac{\langle g, u_j \rangle}{\mu_j} v_j \tag{2.5}$$

where P is the orthogonal projection operator of H_2 onto $\overline{R(K)}$.

Proof First note that, if $g \in \mathcal{D}(K^\dagger) = R(K) + R(K)^\perp$, then $Pg \in R(K)$, and by Theorem 2.3.1,

$$\sum_{j=1}^{\infty} \mu_j^{-2} |\langle Pg, u_j \rangle|^2 < \infty.$$

Also, since $u_j \in \overline{R(K)}$, we have $\langle Pg, u_j \rangle = \langle g, Pu_j \rangle = \langle g, u_j \rangle$. Therefore, the infinite series in equation (2.5) converges in H_1. On the other hand, since $\{v_j\} \subseteq N(K)^\perp$, it follows that the vector defined by

$$v = \sum_{j=1}^{\infty} \mu_j^{-1} \langle Pg, u_j \rangle v_j$$

is also in $N(K)^\perp$, and

$$Kv = \sum_{j=1}^{\infty} \mu_j^{-1} \langle Pg, u_j \rangle K v_j = \sum_{j=1}^{\infty} \langle Pg, u_j \rangle u_j = Pg$$

as $Kv_j = \mu_j u_j$. Therefore v is a least squares solution which lies in $N(K)^\perp = R(K^\dagger)$, i.e. $v = K^\dagger g$. □

The integral operator K in (2.3) has a smoothing effect on u in the sense that high-frequency components in u are smoothed out by the integration. The smoother the kernel $k(\cdot, \cdot)$ is, the faster the singular values μ_j decay to zero [9] and hence

the more ill-posed the problem is, since the high-frequency components will be amplified by μ_j^{-1} in the solution (2.5). Hofmann [10] has defined the severity of ill-posedness: if there exists a positive real number β such that the singular values satisfy $\mu_j = \mathcal{O}(j^{-\beta})$, then the problem is characterized as mildly or moderately ill-posed if $\beta \leq 1$ or $\beta > 1$ respectively. If $\mu_j = \mathcal{O}(e^{-\beta j})$, then the problem is characterized as severely ill-posed. Examples of ill-posed problems of different severities will be given in chapter 5.

Chapter 3

Regularization Theory for First Kind Equations

In this chapter, we will study the general regularization theory as well as some concrete regularization methods for solving Fredholm integral equations of the first kind. In section 3.1, we will review some results in a general class of regularization methods. The "filtering" approach will be used to construct classes of regularization methods. In sections 3.2, 3.3 and 3.4, three of the most important regularization methods: Tikhonov regularization, Landweber iteration and truncated singular value decomposition (TSVD) method will be studied respectively based on the results in section 3.1.

3.1 General regularization theory

Consider the equation

$$Ku = g \tag{3.1}$$

where K is a compact linear operator from a Hilbert space X into a Hilbert space Y. We want to approximate the generalized solution $u^\dagger := K^\dagger g$ for some specific g in the situation that the "exact data" g is unknown, but that only the "noisy

data" g^δ with $\|g^\delta - g\| \leq \delta$ is available for some "noise level" $\delta > 0$. However $K^\dagger g^\delta$ is not a good approximation since, as discussed in the previous chapter, $K^\dagger$ is unbounded in general and as a result $\|K^\dagger g^\delta - K^\dagger g\|$ can be arbitrarily large. Hence we look for some approximation of $u^\dagger = K^\dagger g$, say u_α^δ, which on the one hand, depends continuously on the noisy data g^δ so that it can be computed in a stable way, and on the other hand, has the property that as the noise level δ decreases to zero and the *regularization parameter* α is chosen appropriately, then u_α^δ tends to $u^\dagger$. Generally speaking, a regularization of $K^\dagger$ is a replacement of the unbounded operator $K^\dagger$ by a parameter dependent family of continuous operators $\{R_\alpha\}$. As approximation of $u^\dagger$, we just take $u_\alpha^\delta := R_\alpha g^\delta$, which can then be computed in a stable way (since R_α is assumed to be continuous). A requirement for α is that, if the noise level δ tends to zero, then the regularized solution u_α^δ should tend to $u^\dagger$ [2]. We quote the definition of a regularization method from [2] here.

Definition 3.1.1 *Let $T : X \to Y$ be a bounded linear operator between the Hilbert spaces X and Y, $\alpha_0 \in (0, +\infty]$. For every $\alpha \in (0, \alpha_0)$, let $R_\alpha : Y \to X$ be a continuous (not necessarily linear) operator. The family $\{R_\alpha\}$ is called a regularization or a regularization operator for $T^\dagger$ if, for all $g \in \mathcal{D}(T^\dagger)$ there exists a parameter choice rule $\alpha = \alpha(\delta, g^\delta)$ such that*

$$\lim_{\delta \to 0} \sup\{\|R_{\alpha(\delta, g^\delta)} g^\delta - T^\dagger g\| : g^\delta \in Y, \|g^\delta - g\| \leq \delta\} = 0 \tag{3.2}$$

holds. Here $\alpha : \mathbf{R}^+ \times Y \to (0, \alpha_0)$ is such that

$$\lim_{\delta \to 0} \sup\{\alpha(\delta, g^\delta) : g^\delta \in Y, \|g^\delta - g\| \leq \delta\} = 0. \tag{3.3}$$

For a specific $g \in \mathcal{D}(T^\dagger)$, a pair (R_α, α) is called a regularization method (for solving $Tu = g$) if (3.2) and (3.3) hold.

From the definition, we can see that a regularization method consists of a regularization operator R_α and a parameter choice rule $\alpha(\delta, g^\delta)$. If the regularization

parameter is chosen according to the rule, then the regularized solution $R_{\alpha(\delta,g^\delta)}g^\delta$ will converge to $T^\dagger g$ as the noise level δ tends to zero. The parameter choice rule $\alpha(\delta, g^\delta)$ depends on the noise level δ and the perturbed data g^δ. Following the definition in [2], if the parameter choice rule α only depends on δ but not g^δ, then α is called an *a priori parameter choice rule* and we may just write it as $\alpha = \alpha(\delta)$. Otherwise, α is called an *a posteriori parameter choice rule*. The a priori parameter choice rule may be devised before the actual computation, the a posteriori parameter choice rule depends on the results obtained during the actual computation like the residual $\|Tu_\alpha^\delta - g^\delta\|$ and so may be devised in the computation. In this chapter, we only consider the a priori case. To estimate the error of the computed solution, note that by triangle inequality,

$$\begin{aligned} \|R_\alpha g^\delta - T^\dagger g\| &\leq \|R_\alpha g^\delta - R_\alpha g\| + \|R_\alpha g - T^\dagger g\| \\ &\leq \delta\|R_\alpha\| + \|R_\alpha g - T^\dagger g\|. \end{aligned} \tag{3.4}$$

The first term on the right hand side describes the error of the data multiplied by $\|R_\alpha\|$, which tends to infinity as α tends to zero reflecting the unboundness of $T^\dagger$. The second term is the approximation error $\|R_\alpha g - K^\dagger g\|$ for the exact data, which tends to zero as α tends to zero. It is therefore reasonable that there is an optimal α^* at which the sum of the two terms is minimized. So we want to choose α depending on δ in order to keep the total error as small as possible. Consider $K : X \to Y$ be a compact linear operator with singular system $\{u_j, v_j; \mu_j\}$. Recall that $K^\dagger g$ is given by

$$K^\dagger g = \sum_{j=1}^{\infty} \frac{1}{\mu_j}\langle g, u_j\rangle v_j \tag{3.5}$$

if $g \in \mathcal{D}(K^\dagger)$. Instability arises due to division by small singular values. One way to overcome this instability is to modify the μ_j^{-1}'s by multiplying them by a *regularizing filter function* $q(\alpha, \mu_j)$ for which the product $\mu^{-1}q(\alpha, \mu)$ tends to zero as μ tends to zero [26]. The following theorem, which come from [16] (with some modification in the proof), gives criteria of a filter function.

Theorem 3.1.1 *Let $K : X \to Y$ be a compact linear operator with singular system $\{u_j, v_j; \mu_j\}$ and $q : (0, +\infty) \times (0, \|K\|] \to \mathbf{R}$ be a function with the following properties:*

(1) $|q(\alpha, \mu)| \leq 1$ for all $\alpha > 0$ and $0 < \mu \leq \|K\|$;

(2) For every $\alpha > 0$ there exists $c(\alpha)$ such that

$$|q(\alpha, \mu)| \leq c(\alpha)\mu \qquad \text{for all } 0 < \mu \leq \|K\|;$$

(3a) $\lim_{\alpha \to 0} q(\alpha, \mu) = 1$ for every $0 < \mu \leq \|K\|$.

Then the operator $R_\alpha : Y \to X, \alpha > 0$ defined by

$$R_\alpha g := \sum_{j=1}^{\infty} \frac{q(\alpha, \mu_j)}{\mu_j} \langle g, u_j \rangle v_j, \quad g \in Y$$

is a regularization with $\|R_\alpha\| \leq c(\alpha)$ if $\alpha(\delta) \to 0$ and $\delta c(\alpha(\delta)) \to 0$ as $\delta \to 0$. The function q is called a regularizing filter function for K.

Proof From assumption (2), we have

$$\begin{aligned} \|R_\alpha g\|^2 &= \sum_{j=1}^{\infty} |q(\alpha, \mu_j)|^2 \frac{1}{\mu_j^2} |\langle g, u_j \rangle|^2 \\ &\leq c(\alpha)^2 \sum_{j=1}^{\infty} |\langle g, u_j \rangle|^2 \leq c(\alpha)^2 \|g\|^2 \end{aligned}$$

and so R_α is bounded with $\|R_\alpha\| \leq c(\alpha)$. On the other hand, for $g \in \mathcal{D}(K^\dagger)$, since $\langle g, u_j \rangle = \langle Pg, u_j \rangle = \langle Ku^\dagger, u_j \rangle = \langle u^\dagger, K^* u_j \rangle = \mu_j \langle u^\dagger, v_j \rangle$, we have

$$R_\alpha g = \sum_{j=1}^{\infty} q(\alpha, \mu_j) \langle u^\dagger, v_j \rangle v_j \quad \text{and} \quad K^\dagger g = \sum_{j=1}^{\infty} \langle u^\dagger, v_j \rangle v_j,$$

and hence

$$\|R_\alpha g - K^\dagger g\|^2 = \sum_{j=1}^{\infty} [q(\alpha, \mu_j) - 1]^2 |\langle u^\dagger, v_j \rangle|^2.$$

For any $\epsilon > 0$, there exists $N \in \mathbf{N}$ such that

$$\sum_{n=N+1}^{\infty} |\langle u^\dagger, v_j \rangle|^2 < \frac{\epsilon^2}{8}.$$

By assumption (3a), there exists $\alpha_0 > 0$ such that

$$[q(\alpha, \mu_j) - 1]^2 < \frac{\epsilon^2}{2\|u^\dagger\|^2} \quad \textit{for all } j = 1, 2, \ldots, N \textit{ and } 0 < \alpha \leq \alpha_0.$$

Together with assumption (1), we have

$$\begin{aligned} \|R_\alpha g - K^\dagger g\|^2 &= \sum_{j=1}^{N} [q(\alpha, \mu_j) - 1]^2 |\langle u^\dagger, v_j \rangle|^2 + \sum_{j=N+1}^{\infty} [q(\alpha, \mu_j) - 1]^2 |\langle u^\dagger, v_j \rangle|^2 \\ &< \frac{\epsilon^2}{2\|u^\dagger\|^2} \sum_{j=1}^{N} |\langle u^\dagger, v_j \rangle|^2 + 4 \sum_{j=N+1}^{\infty} |\langle u^\dagger, v_j \rangle|^2 < \frac{\epsilon^2}{2} + \frac{\epsilon^2}{2} = \epsilon^2 \end{aligned}$$

for all $0 < \alpha \leq \alpha_0$ and so $R_\alpha g \to K^\dagger g$ as $\alpha \to 0$. By inequality (3.4),

$$\|R_\alpha g^\delta - K^\dagger g\| \leq \delta c(\alpha) + \|R_\alpha g - K^\dagger g\|.$$

The right hand side tends to zero if $\alpha(\delta) \to 0$ and $\delta c(\alpha(\delta)) \to 0$, hence result follows. □

It can be shown that [16] if (3a) is replaced by the following stronger assumptions:

(3b) There exists $c_1 > 0$ such that

$$|q(\alpha, \mu) - 1| \leq c_1 \frac{\sqrt{\alpha}}{\mu} \quad \textit{for all } \alpha > 0 \textit{ and } 0 < \mu \leq \|K\|$$

and $u^\dagger = K^* z_1 \in R(K^*)$;

(3c) There exists $c_2 > 0$ such that

$$|q(\alpha, \mu) - 1| \leq c_2 \frac{\alpha}{\mu^2} \quad \textit{for all } \alpha > 0 \textit{ and } 0 < \mu \leq \|K\|$$

and $u^\dagger = K^* K z_2 \in R(K^* K)$.

Then we can obtain

$$\|R_\alpha g - K^\dagger g\| \leq c_1 \sqrt{\alpha} \|z_1\| \quad \textit{and} \quad \|R_\alpha g - K^\dagger g\| \leq c_2 \alpha \|z_2\|$$

respectively. In concrete examples, the conditions $u^\dagger \in R(K^*)$ and $u^\dagger \in R(K^* K)$ correspond to smoothness assumptions and boundary conditions on the best approximate solution $u^\dagger$ and are referred to as *source conditions* or *regularity conditions*. The following theorem [16] gives some examples of regularizing filter function.

Theorem 3.1.2 *The following three functions q satisfy the assumptions (1), (2) and (3a-c) respectively:*

(a) $q(\alpha, \mu) = \mu^2/(\alpha + \mu^2)$. *This choice satisfies (2) with* $c(\alpha) = 1/(2\sqrt{\alpha})$. *Assumptions (3b) and (3c) hold with* $c_1 = 1/2$ *and* $c_2 = 1$ *respectively.*

(b) $q(\alpha, \mu) = 1 - (1 - a\mu^2)^{1/\alpha}$ *for some* $0 < a < 2/\|K\|^2$. *In this case, (2) holds with* $c(\alpha) = \sqrt{a/\alpha}$. *Assumptions (3b) and (3c) are satisfied with* $c_1 = 1/\sqrt{2a}$ *and* $c_2 = 1/a$ *respectively.*

(c) Let q be the function defined by

$$q(\alpha, \mu) = \begin{cases} 1, & if\ \mu^2 \geq \alpha, \\ 0, & if\ \mu^2 < \alpha. \end{cases}$$

In this case, (2) holds with $c(\alpha) = 1/\sqrt{\alpha}$. *Assumptions (3b) and (3c) are satisfied with* $c_1 = c_2 = 1$.

Therefore, all of the functions q defined in (a), (b) and (c) are regularizing filter functions. These three filter functions correspond to Tikhonov regularization, Landweber iteration and TSVD method respectively, and will be studied in the next three sections.

3.2 Tikhonov regularization

The idea of a regularization method is to approximate the generalized solution $u^\dagger := K^\dagger g$ in a stable way. However, as the Riemann-Lebesgue lemma shows, certain highly oscillatory noise may be smoothed out by the integral operator, resulting the instability of the solution. Tikhonov's idea ([23], [24], [25]) is to damp out such oscillations and "regularize" the solution by minimizing the following *Tikhonov functional*

$$J_\alpha(u) = \|Ku - g\|^2 + \alpha\|u\|^2 \qquad for\ u \in X$$

where $\alpha > 0$ is the regularization parameter. In this functional, the first term, when small, guarantees that u is "nearly" a least square solution, while the second term, called the *penalty functional* or *regularization functional*, tends to damp out wild instabilities in u [7]. It is easy to see that the quadratic functional J_α has a unique minimizer u_α. For any $w \in X$, let $f(t) = J_\alpha(u_\alpha + tw)$, since $J_\alpha(u)$ is minimum at u_α, the function f should satisfy $f'(0) = 0$, i.e.

$$f'(0) = 2\langle K^*Ku_\alpha - K^*g + \alpha u_\alpha, w\rangle = 0 \qquad \textit{for any } w \in X.$$

Therefore the unique minimizer u_α satifies

$$(K^*K + \alpha I)u_\alpha = K^*g \tag{3.6}$$

which is a regularized form of the normal equation $K^*Ku = K^*g$. Recall that the self-adjoint compact operator K^*K has non-negative eigenvalues. As a result, the operator $K^*K + \alpha I$ has strictly positive eigenvalues for any $\alpha > 0$ and hence has a bounded inverse, so equation (3.6) has a unique solution. Therefore the unique minimizer of $J_\alpha(u)$ is the unique solution of equation (3.6). The solution u_α of equation (3.6) can be written as $u_\alpha = R_\alpha g$ where $R_\alpha := (K^*K + \alpha I)^{-1}K^* : Y \to X$. Then $R_\alpha g$ can be represented as

$$R_\alpha g = \sum_{j=1}^{\infty} \frac{\mu_j}{\alpha + \mu_j^2}\langle g, u_j\rangle v_j = \sum_{j=1}^{\infty} \frac{q(\alpha, \mu_j)}{\mu_j}\langle g, u_j\rangle v_j$$

where $q(\alpha, \mu) = \mu^2/(\alpha + \mu^2)$. The function q is exactly the filter function defined in theorem 3.1.2 (a) and hence R_α defines a regularization for $T^\dagger$ (under certain conditions in part (a) below) and it is called the *Tikhonov regularization*. By applying theorems 3.1.1 and 3.1.2, we obtain the following results [16]:

(a) The operator $R_\alpha := (K^*K + \alpha I)^{-1}K^* : Y \to X$ is a regularization for $T^\dagger$ with $\|R_\alpha\| \le 1/(2\sqrt{\alpha})$ if $\alpha(\delta) \to 0$ and $\delta^2/\alpha(\delta) \to 0$ as $\delta \to 0$.

(b) Let $u^\dagger = K^*z \in R(K^*)$ with $\|z\| \le E$. The choice $\alpha(\delta) = c\delta/E$ for some $c > 0$ leads to the error estimate

$$\|u_\alpha^\delta - u^\dagger\| \le \frac{1}{2}(\frac{1}{\sqrt{c}} + \sqrt{c})\sqrt{\delta E}.$$

(c) Let $u^\dagger = K^*Kz \in R(K^*K)$ with $\|z\| \leq E$. The choice $\alpha(\delta) = c(\delta/E)^{2/3}$ for some $c > 0$ leads to the error estimate

$$\|u_\alpha^\delta - u^\dagger\| \leq (\frac{1}{2\sqrt{c}} + c)E^{\frac{1}{3}}\delta^{\frac{2}{3}}.$$

The regularized solution $R_\alpha g^\delta$ is determined as the unique solution $u_\alpha^\delta \in X$ of the second kind equation $K^*Ku_\alpha^\delta + \alpha u_\alpha^\delta = K^*g^\delta$. From part (a), it can be seen that α needs to be chosen to depend on δ such that α has to converge to zero as δ tends to zero but not as fast as δ^2 does. From parts (b) and (c), we see that the smoother the best approximate solution $u^\dagger$ is, the slower α has to tend to zero. However the convergence can be arbitrarily slow if no a priori assumption about $u^\dagger$ (such as (b) and (c)) is available [21], and the best possible convergence rate obtainable is only $O(\delta^{2/3})$ [2], [16].

3.3 Landweber iteration

The first kind equation (3.1) can also be solved by iterative methods. One of the most well-known iterative methods is the Landweber iteration ([1], [3], [19]) which is defined by

$$u_m = u_{m-1} + a(K^*g - K^*Ku_{m-1}) \qquad for\ m = 1, 2, \ldots \tag{3.7}$$

where $0 < a < 2/\|K\|^2$ is a relaxation parameter. Without loss of generality, we may simply take $u_0 = 0$ [2]. Equation (3.7) is a recursion formula for u_m. By induction, it can be shown that $u_m = R_m g$, where the operator $R_m : Y \to X$ is defined by

$$R_m := a\sum_{k=0}^{m-1}(I - aK^*K)^kK^* \qquad for\ m = 1, 2, \ldots.$$

Then $R_m g$ can be represented as

$$\begin{aligned} R_m g &= a\sum_{k=0}^{m-1}(I - aK^*K)^k K^* \sum_{j=1}^{\infty}\langle g, u_j\rangle u_j \\ &= a\sum_{j=1}^{\infty}\mu_j \sum_{k=0}^{m-1}(1 - a\mu_j^2)^k \langle g, u_j\rangle v_j \\ &= \sum_{j=1}^{\infty}\frac{1}{\mu_j}[1 - (1 - a\mu_j^2)^m]\langle g, u_j\rangle v_j \\ &= \sum_{j=1}^{\infty}\frac{q(m,\mu_j)}{\mu_j}\langle g, u_j\rangle v_j \end{aligned}$$

where $q(m,\mu) = 1 - (1 - a\mu_j^2)^m$. If we take $\alpha = 1/m$, then the function q is the same as the filter function defined in theorem 3.1.2 (b). Therefore, applications of theorems 3.1.1 and 3.1.2 yield the following results [16]:

(a) The operator $R_m := a\sum_{k=0}^{m-1}(I - aK^*K)^k K^* : Y \to X$ defines a regularization for $T^\dagger$ with $\|R_m\| \le \sqrt{am}$ if $m(\delta) \to \infty$ and $\delta^2 m(\delta) \to 0$ as $\delta \to 0$.

(b) Let $u^\dagger = K^*z \in R(K^*)$ with $\|z\| \le E$. The choice $m(\delta)$ with $c_1 E/\delta \le m(\delta) \le c_2 E/\delta$ for some $0 < c_1 < c_2$ leads to the error estimate

$$\|u_m^\delta - u^\dagger\| \le (\sqrt{ac_2} + \frac{1}{\sqrt{2ac_1}})\sqrt{\delta E}.$$

(c) Let $u^\dagger = K^*Kz \in R(K^*K)$ with $\|z\| \le E$. The choice $m(\delta)$ with $c_1(E/\delta)^{2/3} \le m(\delta) \le c_2(E/\delta)^{2/3}$ for some $0 < c_1 < c_2$ leads to the error estimate

$$\|u_m^\delta - u^\dagger\| \le (\sqrt{ac_2} + \frac{1}{ac_1})E^{\frac{1}{3}}\delta^{\frac{2}{3}}.$$

The sequence $u_m^\delta = R_m g^\delta$ is computed by the iteration

$$u_0^\delta = 0; \quad u_m^\delta = u_{m-1}^\delta + a(K^*g^\delta - K^*Ku_{m-1}^\delta) \quad for\ m = 1, 2, \ldots.$$

Similar to Tikhonov's case, the number of iterations m depends on the smoothness of $u^\dagger$ and acts as the regularization parameter. But one thing different from

Tikhonov regularization is that the convergence rate can be better than $O(\delta^{2/3})$. It can be shown that if $u^\dagger \in R((K^*K)^r)$ for some $r \in \mathbf{N}$, the following error estimate [20] can be obtained

$$\|u_m^\delta - u^\dagger\| \leq cE^{\frac{1}{2r+1}}\delta^{\frac{2r}{2r+1}}.$$

3.4 TSVD

Since $K^\dagger g$ is given by equation (3.5) and instability arises due to the small singular values, a straightforward approach to approximate $K^\dagger g$ is to truncate the small singular values. One way to do this is to truncate the singular values μ_j if $\mu_j^2 < \alpha$ for some $\alpha > 0$. The method is called truncated singular value decomposition (TSVD) method. Then the approximation u_α will be given by

$$u_\alpha = \sum_{\mu_j^2 \geq \alpha} \frac{1}{\mu_j}\langle g, u_j\rangle v_j$$

which can be written as $R_\alpha g$ where $R_\alpha : Y \to X$ is defined by

$$R_\alpha g = \sum_{j=1}^{\infty} \frac{q(\alpha, \mu_j)}{\mu_j}\langle g, u_j\rangle v_j, \qquad g \in Y \tag{3.8}$$

where

$$q(\alpha, \mu) = \begin{cases} 1 & if\ \mu^2 \geq \alpha \\ 0 & if\ \mu^2 < \alpha \end{cases}$$

is the filter function defined in theorem 3.1.2 (c). Then by theorems 3.1.1 and 3.1.2 again, we have the following results [16]:

(a) The operator R_α defined by (3.8) is a regularization for $T^\dagger$ with $\|R_\alpha\| \leq 1/\sqrt{\alpha}$ if $\alpha(\delta) \to 0$ and $\delta^2/\alpha(\delta) \to 0$ as $\delta \to 0$.

(b) Let $u^\dagger = K^*z \in R(K^*)$ with $\|z\| \leq E$. The choice $\alpha(\delta) = c\delta/E$ for some $c > 0$ leads to the error estimate

$$\|u_\alpha^\delta - u^\dagger\| \leq (\frac{1}{\sqrt{c}} + \sqrt{c})\sqrt{\delta E}.$$

(c) Let $u^\dagger = K^*Kz \in R(K^*K)$ with $\|z\| \leq E$. The choice $\alpha(\delta) = c(\delta/E)^{2/3}$ for some $c > 0$ leads to the error estimate

$$\|u_\alpha^\delta - u^\dagger\| \leq (\frac{1}{\sqrt{c}} + c)E^{\frac{1}{3}}\delta^{\frac{2}{3}}.$$

As we can see, the applications of Landweber iteration and Tikhonov regularization avoid the knowledge of the singular system which is an advantage over TSVD method.

Chapter 4

Multilevel Algorithms for Ill-posed Problems

The theory of multigrid methods for finite difference or finite element applications to differential equations is well-developed. However, for integral equations, the situation is quite different. There have been few papers on multigrid techniques for Fredholm integral equations of the first kind. The main advantage of multigrid methods over traditional iterative methods like Landweber iteration and conjugate gradient method is that multigrid methods save a lot of computational effort in the coarse grid while yielding a fast convergence by effective preconditioning. In this chapter, we will present the multilevel algorithms proposed by King which can be found in [14] and the results therein. Section 4.1 outlines the discretization and defines some norms, projections and operators that will be used in the analysis in section 4.2, in which the multilevel algorithms are defined. The applications of the multilevel operators will be presented in section 4.3, and numerical examples using the multilevel algorithms will be given in section 5.4.

4.1 Basic assumptions and definitions

Consider a compact operator $K : H_1 \to H_2$ where H_1 and H_2 are real Hilbert spaces. Our problem is to solve

$$Ku = g$$

where the data g is known imprecisely, i.e. we are only given g^δ with $\|g - g^\delta\| \leq \delta$, where δ is the noise level. In this section we will first define the subspaces in which the approximate solution is obtained. Then we will define some multilevel operators and give some of their properties that will be used in the next section.

Suppose there are nested finite dimensional subspaces of $\overline{R(K)} = N(K^*)^\perp$, say $W_1 \subset W_2 \subset \cdots W_m \subset N(K^*)^\perp$. Assume $N(K^*) = \{0\}$, so the subspaces W_i are easy to construct. We denote the inner product on H_1 and H_2 by $(\cdot, \cdot)$ and $\langle \cdot, \cdot \rangle$ respectively, and assume there is an inner product $\langle \cdot, \cdot \rangle_j$ defined on $W_j \times W_j$ that is equivalent to $\langle \cdot, \cdot \rangle$, i.e. there exist $0 < \alpha_j \leq \beta_j$ such that

$$\alpha_j ||w||_j^2 \leq ||w||^2 \leq \beta_j ||w||_j^2 \qquad \forall w \in W_j \tag{4.1}$$

where $\|\cdot\|$ and $\|\cdot\|_j$ denote the norm induced by $\langle \cdot, \cdot \rangle$ and $\langle \cdot, \cdot \rangle_j$ respectively. On $W_m \times W_m$, we define the symmetric bilinear form

$$a(u, v) = (K^*u, K^*v) + \lambda \langle u, v \rangle_m \qquad \forall u, v \in W_m$$

where $\lambda \geq 0$ is a regularization parameter, and we define $|||\cdot|||$ as the norm induced by $a(\cdot, \cdot)$, i.e.

$$|||u||| = a^{1/2}(u, u) \qquad \forall u \in W_m.$$

In application, we may consider the inner products $\langle \cdot, \cdot \rangle$ and $\langle \cdot, \cdot \rangle_m$ correspond to L_2 and l_2 respectively. We define the following projection operators on the subspaces W_j. Let $P_j : H_2 \to W_j$ and $Q_j : W_{j+1} \to W_j$ be defined by

$$\begin{aligned} a(P_j u, v) &= a(u, v) \qquad \forall v \in W_j, \\ \langle Q_j u, v \rangle_{j+1} &= \langle u, v \rangle_{j+1} \qquad \forall v \in W_j. \end{aligned}$$

We also define the restriction operator $R_j : W_{j+1} \to W_j$ by

$$\langle R_j w, v\rangle_j = \langle w, v\rangle_{j+1} \quad \forall v \in W_j.$$

Moreover we define $A_j : W_j \to W_j$, the discrete versions of $KK^* + \lambda I$, by

$$\langle A_j u, v\rangle_j = a(u, v) \quad \forall v \in W_j.$$

It can be seen that A_j is a symmetric operator in the inner products $\langle\cdot,\cdot\rangle_j$ and $a(\cdot,\cdot)$, and $R_j A_{j+1} = A_j P_j$ for $j = 1, 2, \ldots, m-1$. In addition, we set

$$\gamma_j := ||K^*(I - Q_j)|| = \sup_{w \in W_{j+1}} \frac{||K^*(I - Q_j)w||}{||w||_{j+1}}$$

and assume that $\gamma_j < 1$ for each j. We also set

$$\mu_j := \min_{w \in W_j} \frac{||K^* w||^2}{||w||_j^2}$$

which is the minimal eigenvalue of the discrete version of $\hat{K} := KK^*$ on W_j. It is easy to show that $\mu_j \leq \gamma_{j-1}^2$ for $j = 2, 3, \ldots, m$. We define the operator $T_j : W_j \to W_j$ by

$$T_j := I - \xi_j (I - Q_{j-1}) A_j$$

where $\xi_j > 0$ will be specified later. This operator will play an important role in the multilevel analysis in the next section. The following lemma gives a spectral property of T_j.

Lemma 4.1.1 *The operator T_j is symmetric in $a(\cdot,\cdot)$ and the spectrum of T_j, $\sigma(T_j)$, satisfies $\sigma(T_j) \subset (0, 1]$ if*

$$0 < \xi_j < (\alpha_m^{-1} \beta_j \lambda + \gamma_{j-1}^2)^{-1}. \tag{4.2}$$

On the other hand, it is straightforward to show that

$$\langle A_j w, w\rangle_j \geq (\mu_j + \beta_m^{-1} \alpha_j \lambda) ||w||_j^2 \quad \forall w \in W_j. \tag{4.3}$$

The following lemma shows that T_j is a reducer in $a(\cdot,\cdot)$ on the orthogonal complement $W_{j-1}^{\perp} = \{w \in W_j : a(w, v) = 0 \ \ \forall v \in W_{j-1}\}$.

Lemma 4.1.2 *For $w \in W_j$ and ξ_j specified in (4.2),*

$$0 < a(T_j(I - P_{j-1})w, (I - P_{j-1})w) \leq \delta_j |||(I - P_{j-1})w|||^2,$$

where $\delta_j = 1 - (\mu_j + \beta_m^{-1}\alpha_j\lambda)\xi_j < 1$.

Proof Since $P_{j-1}Q_{j-1} = Q_{j-1}$, it follows that $(I - P_{j-1})Q_{j-1} = 0$ and so

$$(I - P_{j-1})T_j = (I - P_{j-1})(I - \xi_j A_j).$$

Since (4.3) is equivalent to

$$a(w, w) \geq (\mu_j + \beta_m^{-1}\alpha_j\lambda)||w||_j^2 \qquad \forall w \in W_j,$$

we have for $w \in W_j$,

$$\begin{aligned}
& a(T_j(I - P_{j-1})w, (I - P_{j-1})w) \\
= \; & a((I - \xi_j A_j)(I - P_{j-1})w, (I - P_{j-1})w) \\
= \; & |||(I - P_{j-1})w|||^2 - \xi_j a(A_j^{1/2}(I - P_{j-1})w, A_j^{1/2}(I - P_{j-1})w) \\
\leq \; & |||(I - P_{j-1})w|||^2 - \xi_j(\mu_j + \beta_m^{-1}\alpha_j\lambda)||A_j^{1/2}(I - P_{j-1})w||_j^2 \\
= \; & [1 - (\mu_j + \beta_m^{-1}\alpha_j\lambda)\xi_j]|||(I - P_{j-1})w|||^2.
\end{aligned}$$

The other inequality follows from lemma 4.1.1. □

The reduction factor, δ_j, depends on the minimal eigenvalue μ_j and the regularization parameter λ. It is desirable to have the reduction factor as small as possible. From lemma 4.1.1, one reasonable choice of ξ_j is $(\alpha_m^{-1}\beta_j\lambda + 2\gamma_{j-1}^2)^{-1}$ and with this choice, our reduction factor will be $\delta_j = 1 - (\mu_j + \beta_m^{-1}\alpha_j\lambda)(\alpha_m^{-1}\beta_j\lambda + 2\gamma_{j-1}^2)^{-1}$.

4.2 Multilevel analysis

In this section, two symmetric multilevel operators will be defined and the reduction properties of these two operators will be analyzed. Define the symmetric multilevel operator $B_j : W_j \to W_j$ inductively as follows:

Algorithm 1 Set $B_1 = A_1^{-1}$ and suppose B_k has been defined for $k \leq j-1$. For $f \in W_j$ define $B_j f = w_1 + w_2$ where

(i) $w_0 = B_{j-1} R_{j-1} f$,

(ii) $w_1 = w_0 + \xi_j (I - Q_{j-1})(f - A_j w_0)$,

(iii) $w_2 = B_{j-1} R_{j-1} (f - A_j w_1)$,

where ξ_j is specified in (4.2).

Then the following identity can be obtained

$$I - B_j A_j = S_j T_j S_j$$

where $S_j = I - P_{j-1} + (I - B_{j-1} A_{j-1}) P_{j-1}$. By using the Cauchy-Schwarz inequality for $a(T_j \cdot, \cdot)$ and lemma 4.1.2, it can be shown that

$$|||(I - P_{j-1}) T_j v|||^2 \leq \delta_j a(T_j v, v) \qquad \forall v \in W_j. \tag{4.4}$$

The following theorem shows that $I - B_j A_j$ is a reducer in the $a(\cdot, \cdot)$ inner product.

Theorem 4.2.1 *There exists* $\varepsilon_j \in (0, 1)$ *such that*

$$0 \leq a((I - B_j A_j) w, w) \leq \varepsilon_j |||w|||^2 \qquad \forall w \in W_j$$

and $\varepsilon_j = \varepsilon_{j-1}^2 + (1 - \varepsilon_{j-1}^2) \delta_j$.

Proof By Cauchy-Schwarz inequality, $a((I - B_j A_j) w, w) \leq |||(I - B_j A_j) w||| \, |||w|||$, so it suffices to show that

$$|||(I - B_j A_j) w||| \leq \varepsilon_j |||w||| \qquad \forall w \in W_j. \tag{4.5}$$

Inequality (4.5) can be proved by using induction. It is trivial for $j = 1$ since $B_1 = A_1^{-1}$. Suppose (4.5) is true for $j-1$. Since $S_j = I - P_{j-1} + (I - B_{j-1} A_{j-1}) P_{j-1}$ and by orthogonality,

$$|||(I - B_j A_j) w|||^2 = |||(I - P_{j-1}) T_j S_j w|||^2 + |||(I - B_{j-1} A_{j-1}) P_{j-1} T_j S_j w|||^2.$$

By using the induction hypothesis and Pythagorean theorem,

$$\begin{aligned} |||(I - B_j A_j)w|||^2 &\leq |||(I - P_{j-1})T_j S_j w|||^2 + \varepsilon_{j-1}^2 |||P_{j-1} T_j S_j w|||^2 \\ &= (1 - \varepsilon_{j-1}^2)|||(I - P_{j-1})T_j S_j w|||^2 + \varepsilon_{j-1}^2 |||T_j S_j w|||^2. \end{aligned}$$

Applying (4.4) and lemma 4.1.1 gives

$$\begin{aligned} |||(I - B_j A_j)w|||^2 &\leq (1 - \varepsilon_{j-1}^2)\delta_j a(T_j S_j w, S_j w) + \varepsilon_{j-1}^2 a(T_j S_j w, S_j w) \\ &\leq [(1 - \varepsilon_{j-1}^2)\delta_j + \varepsilon_{j-1}^2] a((I - B_j A_j)w, w) \\ &\leq \varepsilon_j |||(I - B_j A_j)w||| \; |||w||| \end{aligned}$$

and hence result follows. □

Let $\hat{\delta} = \max\{\delta_j : 2 \leq j \leq m\}$. The following corollary shows that the reduction factor is independent of the level.

Corollary 4.2.2 *For* $1 \leq j \leq m$,

$$0 \leq a((I - B_j A_j)w, w) \leq \hat{\varepsilon} |||w|||^2$$

where $\hat{\varepsilon} = 1 - (1 - \hat{\delta})(1 - \hat{\delta}^2)^{m-2}$.

Next define another symmetric multilevel operator $C_j : W_j \to W_j$ which will be used in the numerical experiments in section 5.4.

Algorithm 2 Set $C_1 = A_1^{-1}$ and suppose C_k has been defined for $k \leq j - 1$. For $f \in W_j$ define $C_j f = w_4$ where

(i) $w_1 = \xi_j (I - Q_{j-1}) f$,

(ii) for $k = 1, 2$

$$w_{k+1} = w_k + C_{j-1} R_{j-1}(f - A_j w_k),$$

(iii) $w_4 = w_3 + \xi_j (I - Q_{j-1})(f - A_j w_3)$,

where ξ_j is specified in (4.2).

Algorithm 2 is obtained by interchanging the order of operations in algorithm 1. Step (ii) is the coarse grid correction, while steps (i) and (iii) correspond to the pre-smoothing and post-smoothing operations respectively. The smoothing operator used in both steps (i) and (iii) is given by $\xi_j(I - Q_{j-1})$, which acts as a stabilized rough approximation of the inverse of A_j on the "high-frequency part" of f, while just removing the "low-frequency part" $Q_{j-1}f$ [11]. For this operator C_j, similar to the previous one, the following identity can be obtained

$$I - C_j A_j = T_j L_j T_j$$

where $L_j = I - P_{j-1} + (I - C_{j-1}A_{j-1})^2 P_{j-1}$. The following theorem shows that $I - C_j A_j$ is also a reducer in the $a(\cdot,\cdot)$ inner product and has the same reduction as $I - B_j A_j$.

Theorem 4.2.3 *For $w \in W_j$,*

$$0 \leq a((I - C_j A_j)w, w) \leq \varepsilon_j a(w, w). \tag{4.6}$$

where ε_j is the same as in theorem 4.2.1.

Proof Again (4.6) can be proved by induction. Assume (4.6) is true for $j-1$. Since $L_j = I - P_{j-1} + (I - C_{j-1}A_{j-1})^2 P_{j-1}$, using the induction hypothesis, it follows that

$$\begin{aligned}
a((I - C_j A_j)w, w) &= a(L_j T_j w, T_j w) \\
&= a((I - P_{j-1})T_j w, T_j w) + a((I - C_{j-1}A_{j-1})^2 P_{j-1} T_j w, T_j w) \\
&\leq a((I - P_{j-1})T_j w, T_j w) + \varepsilon_{j-1}^2 a(P_{j-1} T_j w, T_j w) \\
&= |||(I - P_{j-1})T_j w|||^2 + \varepsilon_{j-1}^2 |||P_{j-1} T_j w|||^2 \\
&= (1 - \varepsilon_{j-1}^2)|||(I - P_{j-1})T_j w|||^2 + \varepsilon_{j-1}^2 |||T_j w|||^2.
\end{aligned}$$

The last equality is due to Pythagorean theorem. Finally applying (4.4) and lemma 4.1.1, we get

$$\begin{aligned} a((I - C_j A_j)w, w) &\leq (1 - \varepsilon_{j-1}^2)\delta_j a(T_j w, w) + \varepsilon_{j-1}^2 a(T_j w, w) \\ &\leq \varepsilon_j a(w, w). \end{aligned}$$

□

One gets the same bound as in corollary 4.2.2 for $I - C_j A_j$ and hence these two multilevel operators are essentially equivalent in terms of reduction in $a(\cdot, \cdot)$. One can also get a non-symmetric multilevel operator by omitting the post-smoothing step (iii) in algorithm 2. However the reduction obtained for this non-symmetric operator is only the square root of that for the symmetric ones.

4.3 Applications

Consider problems of the form

$$A_m w_{m,\lambda} = \hat{g} \tag{4.7}$$

for some given $\hat{g} \in W_m$. The operator A_m is some discrete version of $KK^* + \lambda I$. The aim is to approximate $u_{m,\lambda} := K^* w_{m,\lambda}$ by some iterative methods that uses the multilevel operators defined before. In this section, the application of the multilevel operators in Landweber iterative method and conjugate gradient (CG) method will be introduced, and the error estimates for these two methods will be given. The full multilevel algorithm for the solution of (4.7) will be defined at the end.

4.3.1 Preconditioned iterative methods with nonzero regularization parameter

One can use B_m to precondition the Landweber iteration:

$$w^k = w^{k-1} + B_m(\hat{g} - A_m w^{k-1}).$$

The approximate solution is then taken as $u^k := K^* w^k$. So the error in the k-th iterate is $e^k := u_{m,\lambda} - u^k = K^* \eta^k$ where $\eta^k := w_{m,\lambda} - w^k \in W_m$. Moreover $\eta^k = (I - B_m A_m)\eta^{k-1}$, hence by corollary 4.2.2,

$$\begin{aligned} ||e^k|| &\leq |||\eta^k||| = |||(I - B_m A_m)\eta^{k-1}||| \\ &\leq \hat{\varepsilon}|||\eta^{k-1}||| \leq \hat{\varepsilon}^k|||\eta^0|||. \end{aligned}$$

Beside Landweber iterative method, one can also apply B_m in a preconditioned CG method:

$$B_m A_m w_{m,\lambda} = B_m \hat{g}.$$

The average reduction per step is [12]

$$\rho = \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1}$$

where $\kappa = \kappa(B_m A_m)$ is the condition number of $B_m A_m$. From corollary 4.2.2,

$$(1 - \hat{\varepsilon})a(w, w) \leq a(B_m A_m w, w) \leq a(w, w),$$

which implies $\kappa(B_m A_m) \leq 1/(1 - \hat{\varepsilon})$, and hence the reduction per step for this method is

$$\rho = \frac{1 - \sqrt{1 - \hat{\varepsilon}}}{1 + \sqrt{1 - \hat{\varepsilon}}}.$$

4.3.2 Preconditioned iterative methods with zero regularization parameter

One can also use B_m to precondition an iterative method for $\lambda = 0$, i.e. use B_m as defined previously with $\lambda \neq 0$ to precondition the iterative solution of the

problem: Find $w_m \in W_m$ such that

$$(K^* w_m, K^* v) = \langle \hat{g}, v \rangle_m \quad \forall v \in W_m.$$

To be more precise, let $\hat{K}_j : W_j \to W_j$ be defined by

$$(K^* w, K^* v) = \langle \hat{K}_j w, v \rangle_j \quad \forall v \in W_j.$$

Then $w_m = \hat{K}_m^{-1} \hat{g}$. To use B_m as a preconditioner in the preconditioned CG method applied to $\hat{K}_m w_m = \hat{g}$, it is necessary to estimate the condition number $\kappa(B_m \hat{K}_m)$. It can be shown that for any $w \in W_m$,

$$(1 - \hat{\varepsilon})\theta^{-1} \langle \hat{K}_m w, w \rangle_m \leq \langle B_m \hat{K}_m w, \hat{K}_m w \rangle_m \leq \langle \hat{K}_m w, w \rangle_m \tag{4.8}$$

which gives the condition number bound

$$\kappa(B_m \hat{K}_m) \leq \theta / (1 - \hat{\varepsilon})$$

where $\theta = 1 + \lambda \mu_m^{-1}$. For Landweber iteration, let $\mathcal{A}$ denote the symmetric positive definite bilinear form on $W_m \times W_m$ defined by

$$\mathcal{A}(u, v) = (K^* u, K^* v) \qquad \forall u, v \in W_m$$

and $|||\cdot|||_{\mathcal{A}}$ denote the norm induced by $\mathcal{A}(\cdot, \cdot)$. Then by (4.8),

$$(1 - \hat{\varepsilon})\theta^{-1} \mathcal{A}(w, w) \leq \mathcal{A}(B_m \hat{K}_m w, w) \leq \mathcal{A}(w, w)$$

which gives

$$0 \leq \mathcal{A}((I - B_m \hat{K}_m) w, w) \leq \rho \mathcal{A}(w, w)$$

where $\rho = 1 - (1 - \hat{\varepsilon})\theta^{-1}$. It follows that the Landweber iteration $w^k = w^{k-1} + B_m(\hat{g} - \hat{K}_m w^{k-1})$ satisfies

$$\begin{aligned} ||e^k|| &= |||\eta^k|||_{\mathcal{A}} = |||(I - B_m \hat{K}_m)\eta^{k-1}|||_{\mathcal{A}} \\ &\leq \rho |||\eta^{k-1}|||_{\mathcal{A}} \\ &\leq \rho^k |||\eta^0|||_{\mathcal{A}} = \rho^k ||e^0||. \end{aligned}$$

4.3.3 Full multilevel algorithm

The full multilevel algorithm for the solution of $A_m w_{m,\lambda} = \hat{g}$ is defined as follows:

Algorithm 3 For any integer $n \geq 1$, let $\{w_j^n\} \subset W_j$ be defined as:

(i) for j=1, set $w_1^n = A_1^{-1}\hat{g}_1$,

(ii) for $j > 1$, compute n iterations of the j-level iteration
$w_j^k = w_j^{k-1} + B_j(\hat{g}_j - A_j w_j^{k-1}), \quad k \geq 1$

with initial guess $w_j^0 = w_{j-1}^n$, where $\hat{g}_j = R_j \cdots R_{m-1}\hat{g}$ and n is suitably chosen.

Then $u_m^n := K^* w_m^n$ is taken as the approximate solution. For the choice $n \geq -3\ln 2/\ln\hat{\varepsilon}$, it can be shown that there results the quasioptimal estimate

$$||e_m^n|| \leq Ch_m^2||z||$$

where the condition $u^\dagger = K^* z$ for some z is assumed.

Chapter 5

Numerical Experiments

In this chapter, we will investigate the performance of the numerical methods that have been mentioned in chapters 3 and 4. We will study integral equations and differential equations in sections 5.1 and 5.2 respectively. Discretization of the two kinds of equations as well as singular values and singular vectors of the discretized systems will be shown. In section 5.3, we will see the performance of three classical regularization methods: Tikhonov regularization, TSVD method and Landweber iteration, and the effect of parameters on their convergence through numerical experiments of three test problems. The performance of multilevel methods will be studied in section 5.4, in which general convergence and the effect of multilevel parameters on convergence will be investigated.

5.1 Integral equations

In this section, we will first show the discretization of Fredholm integral equations of the first kind, followed by three test problems which come from first kind Fredholm integral equations. Then we will study the singular values, singular vectors and condition numbers of the discretized systems, and the effect of condition numbers on numerical accuracies in solving linear systems.

5.1.1 Discretization

Since our problem is infinite dimensional, it needs to be discretized into a finite dimensional one before the application of computer algorithms. Consider the Fredholm integral equation of the first kind

$$\int_0^1 k(s,t)u(t)\,dt = g(s), \quad 0 \le s \le 1.$$

Applying the collocation method at the collocation points $\{s_i : 1 \le i \le M\}$, we have

$$\int_0^1 k(s_i,t)u(t)\,dt = g(s_i), \quad 1 \le i \le M.$$

Then we can approximate the integral by some quadrature rule

$$\int_0^1 \phi(t)\,dt \approx \sum_{j=1}^{N} w_j \phi(t_j)$$

(in the experiments, midpoint rule is used, and we always set $M{=}N$ and $s_i = i/N$), and a system of linear equations is resulted

$$\sum_{j=1}^{N} w_j k(s_i,t_j)u(t_j) = g(s_i), \quad i = 1,2,\ldots,M.$$

We denote this system by

$$\mathbf{Ku} = \mathbf{g}$$

where

$$\mathbf{K}_{ij} = w_j k(s_i,t_j), \quad \mathbf{u}_j = u(t_j), \quad \mathbf{g}_i = g(s_i), \quad 1 \le i \le M,\ 1 \le j \le N.$$

In practice, measurement error always exists, we let $\mathbf{e} \in \mathbf{R}^N$ be the measurement noise (normally distributed with mean 0) in our data such that the noise-to-signal ratio is given by

$$\frac{\|\mathbf{e}\|_2}{\|\mathbf{g}\|_2} = \delta.$$

The system we actually solve is

$$\mathbf{Ku}^\delta = \mathbf{g}^\delta, \tag{5.1}$$

where

$$\mathbf{g}^\delta = \mathbf{g} + \mathbf{e}$$

called the analytic data, or

$$\mathbf{g}^\delta = \mathbf{K}\mathbf{u} + \mathbf{e}$$

called the synthetic data. Discrete ill-posed problems arise from the discretization of Fredholm integral equations of the first kind and the system (5.1) will be highly ill-conditioned. The singular values of the matrix $\mathbf{K}$ will decay gradually to zero [9]. All these properties will be observed in the numerical experiments later.

5.1.2 Test problems

Our aim is to solve the integral equation

$$\int_0^1 k(s,t)u(t)\,dt = g(s), \quad 0 \le s \le 1 \tag{5.2}$$

numerically. In this subsection, we will pose three test problems which are the form of (5.2) with different kernels and data functions. These test problems will be solved numerically by different methods in the numerical experiments later.

Test problem 1 This is a classical example of an ill-posed problem and arises from physical sciences (see [7]). Consider the model equation (5.2) where the kernel and the data functions are

$$k(s,t) = \begin{cases} t(1-s) & \text{if } 0 \le t \le s \\ s(1-t) & \text{if } s \le t \le 1 \end{cases} \quad \textit{and} \quad g(s) = -\frac{1}{\pi^2}\sin \pi s$$

respectively. The exact solution is given by $u(t) = -\sin \pi t$. Since the kernel function is square integrable over $[0,1] \times [0,1]$, the integral operator is compact. The symmetric kernel $k(\cdot,\cdot)$ is not differentiable across the line $s = t$. The singular

values and functions for $i = 1, 2, \ldots$ are given by (see [9, 17])

$$\begin{aligned} \mu_i &= (i\pi)^{-2}, \\ u_i(s) &= \pm\sqrt{2}\sin(i\pi s), \\ v_i(t) &= \mp\sqrt{2}\sin(i\pi t). \end{aligned}$$

Since the singular values are proportional to i^{-2}, according to the definition by Hofmann, the problem is characterized as moderately ill-posed.

Test problem 2 Consider the model problem (5.2) with

$$k(s,t) = \cos st\ ; \quad g(s) = \frac{s^2-2}{s^3}\sin s + \frac{2}{s^2}\cos s.$$

The exact solution is given by $u(t) = t^2$. This kernel is smoother than the one in the last problem , so it is expected that this problem is more ill-posed than the previous one and this is reflected in the decay of the singular values as we will see later.

Test problem 3 This test problem is different from the previous two problems in that the solution is discontinuous. We consider a discontinuous solution

$$u(t) = \begin{cases} 1 & \text{for } 1/3 < t < 2/3, \\ 0 & \text{elsewhere.} \end{cases}$$

Here our kernel and the data function are given by

$$k(s,t) = te^{-st^2} \quad \textit{and} \quad g(s) = \frac{e^{-s/9} - e^{-4s/9}}{2s}$$

respectively (see [9]). Again this kernel is very smooth, the singular values decay very fast and the problem is severely ill-posed, which will be seen in the next subsection.

5.1.3 Singular values, singular vectors and condition numbers

In this subsection, we will study the singular values, singular vectors and condition numbers of the matrix $\mathbf{K}$ formed from the discretization of test problems 1, 2 and 3. Set $N = 40$. We denote σ_i the singular values of the matrix $\mathbf{K}$. Figures 5.1, 5.2 and 5.3 show the distribution of the singular values and some of the corresponding right singular vectors for test problems 1, 2 and 3 respectively. From figure 5.1, we see that σ_i decreases like $i^{-\beta}$ for some $\beta > 0$, which agrees with the moderately ill-posedness of test problem 1. We observe that the singular vector $\mathbf{v_i}$ becomes more oscillating as the singular value σ_i decreases. The odd singular vectors are symmetric along the line $t = 0.5$ and the even singular vectors are symmetric about the point $(0.5, 0)$. The smallest singular value of the matrix $\mathbf{K}$, for any N, is zero by our discretization. From figure 5.2, we see that the singular value decreases exponentially at the beginning and then levels off due to limitations of the arithmetic accuracy of the computer. Hence test problem 2 is characterized as a severely ill-posed problem. Similar to test problem 1, the singular vector becomes more oscillating as the singular value decreases. The singular vectors starting from $\mathbf{v_7}$ are inaccurately computed due to the limited accuracy of the computer. The results in figure 5.3 is similar to that in figure 5.2 and the singular vectors are inaccurately computed starting from $\mathbf{v_{10}}$. Table 5.1 shows the condition numbers of $\mathbf{K}$ when different grid numbers N are used for test problems 1, 2 and 3. The condition number is defined to be the ratio of the largest positive singular value to the smallest positive singular value. We observe that the condition number increases (generally) as the grid number increases, and hence the more ill-conditioned the matrix $\mathbf{K}$ is, reflecting the ill-posedness of the underlying problem.

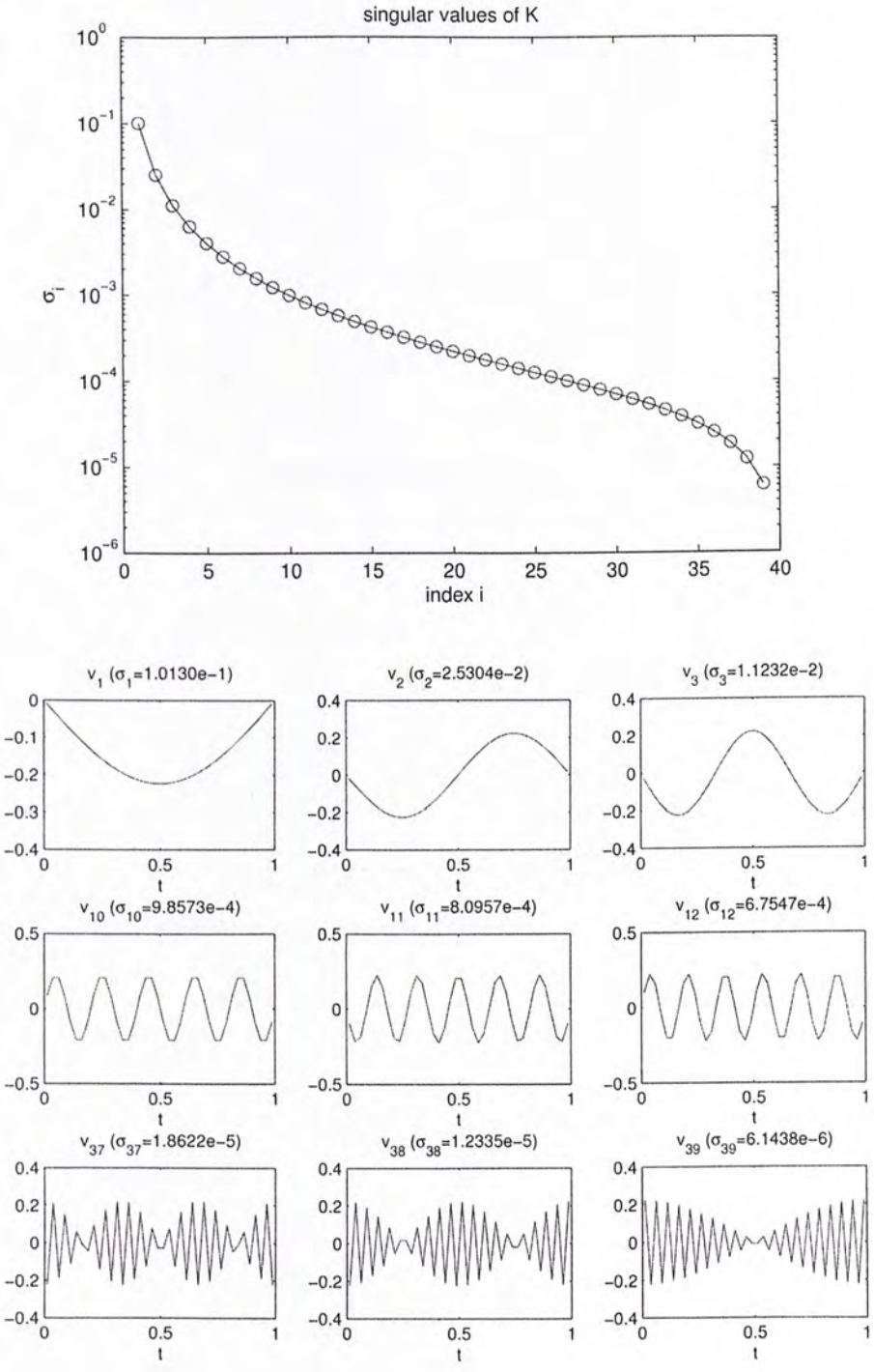

Figure 5.1: Test problem 1: the top subplot shows the distribution of the singular values σ_i of the matrix $\mathbf{K}$, and the subplots below it show the corresponding right singular vectors $\mathbf{v_i}$ for $i = 1, 2, 3, 10, 11, 12, 37, 38, 39$.

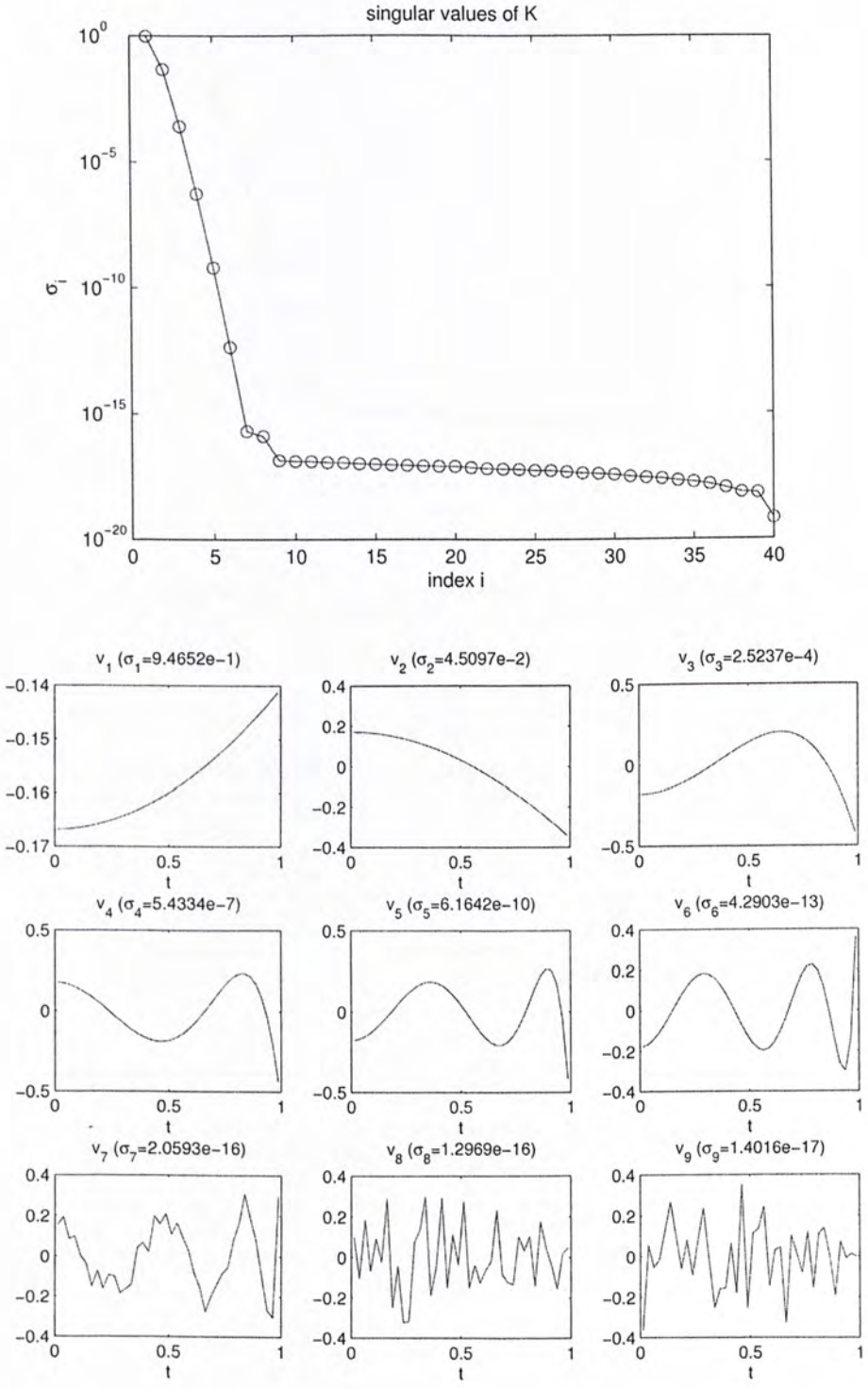

Figure 5.2: Test problem 2: the top subplot shows the distribution of the singular values σ_i of the matrix $\mathbf{K}$, and the subplots below it show the corresponding right singular vectors $\mathbf{v_i}$ for $i = 1, 2, \ldots, 9$.

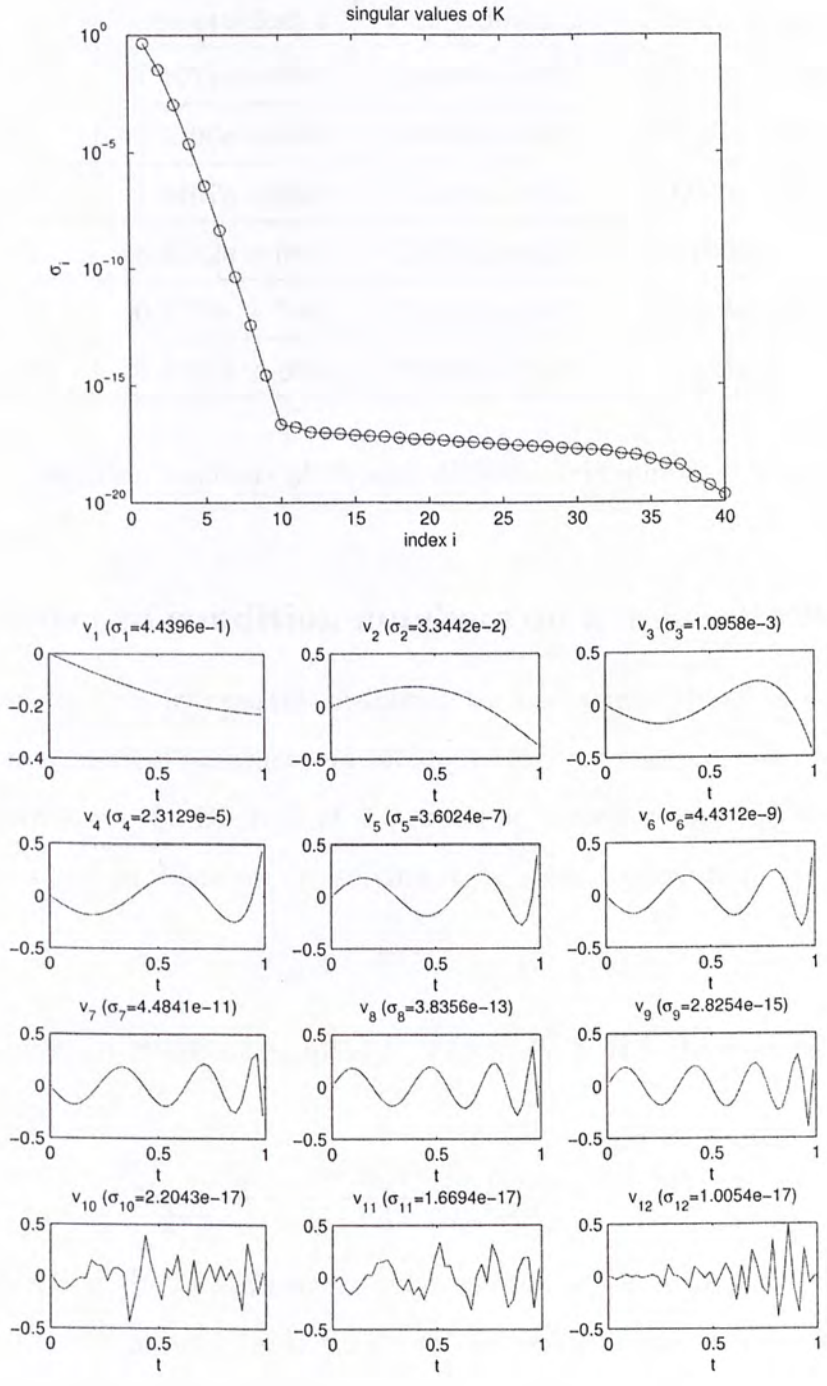

Figure 5.3: Test problem 3: the top subplot shows the distribution of the singular values σ_i of the matrix $\mathbf{K}$, and the subplots below it show the corresponding right singular vectors $\mathbf{v_i}$ for $i = 1, 2, \ldots, 12$.

N	Test problem 1	Test problem 2	Test problem 3
4	$1.4071e+001$	$5.0848e+006$	$1.3467e+005$
8	$1.2706e+002$	$1.4539e+017$	$2.0520e+014$
16	$1.0467e+003$	$1.1391e+019$	$3.6640e+018$
32	$8.4342e+003$	$2.8316e+018$	$7.2708e+018$
64	$6.7596e+004$	$1.0261e+019$	$8.2994e+019$
128	$5.4101e+005$	$3.3125e+019$	$4.1284e+020$

Table 5.1: Condition numbers of $\mathbf{K}$ with different grid numbers N for test problems 1, 2 and 3.

5.1.4 Effect of condition numbers on numerical accuracies

In this subsection, we will use test problem 2 to demonstrate the effect of condition numbers on numerical accuracies in solving linear systems. Consider solving the linear system $Ax = b$. Because of the rounding errors, or the observation data errors, the actual problem we are solving is the perturbed system

$$(A + E)\tilde{x} = b + e.$$

Denote the 2-norm condition number of A by $\kappa(A)$. It is well-known in numerical linear algebra that

$$\frac{\|\tilde{x} - x\|_2}{\|x\|_2} \leq \frac{\kappa(A)}{1 - \|A^{-1}E\|_2}\Big(\frac{\|e\|_2}{\|b\|_2} + \frac{\|E\|_2}{\|A\|_2}\Big),$$

which means that the relative error of the solution is approximately proportional to the condition number. Table 5.2 shows the condition numbers of $\mathbf{K}$, relative residuals and relative errors for solving $\mathbf{K}x = b$, where the matrix $\mathbf{K}$ is generated by the kernel in test problem 2, when different grid numbers N are used. The fifth column is the product of condition number of $\mathbf{K}$ and relative residual. *The exact solution x is a random vector in $\mathbf{R}^N$ whose elements are uniformly distributed in the interval $(0, 1)$. The vector b is formed by the multiplication $b = \boldsymbol{K}x$, and $\tilde{x}$ is*

our computed solution by Gaussian elimination with partial pivoting. Naively, it appears that a small residual would imply that the computed solution $\tilde{x}$ is close to the exact solution x. However this is only true if the condition number of A is not large since

$$\frac{\|\tilde{x} - x\|_2}{\|x\|_2} \leq \kappa(A) \frac{\|b - A\tilde{x}\|_2}{\|b\|_2}. \tag{5.3}$$

That means even the relative residual is very small, the relative error can be large if the condition number is large. This phenomenon is observed in table 5.2 due to the large condition number of $\mathbf{K}$ resulting from the ill-posedness of the underlying problem. We also see that the fourth column is always smaller than the fifth one, verifying inequality (5.3). The numerical accuracies are greatly affected by the condition numbers in solving linear system of equations. In this example, the machine accuracy is $\varepsilon_M = 2.2204e - 016$. From the last column, which is the infinity norm of $\tilde{x} - x$, we observe that the computed solution roughly loses k digits in solving the system $\mathbf{K}x = b$ if $\kappa(\mathbf{K}) \approx 10^k$. We cannot conclude that the computed solution is a good approximation to the exact solution even the relative residual is small when solving ill-conditioned systems. Regularization methods are needed to be imposed on such systems to stabilize the solutions. Figure 5.4 shows the relative errors and relative residuals when different grid numbers N are used. We see that there is a large gap between the relative error and relative residual for test problem 2.

5.2 Differential equations

In this section, we will first show the discretization of a boundary value problem using finite element method. Then we will study the singular values, singular vectors and condition numbers of the discretized system, and compare them with those in the integral equation case in subsection 5.1.3.

N	$\kappa(\mathbf{K})$	$\frac{\|b-\mathbf{K}\tilde{x}\|_2}{\|b\|_2}$	$\frac{\|\tilde{x}-x\|_2}{\|x\|_2}$	$\kappa(\mathbf{K})\frac{\|b-\mathbf{K}\tilde{x}\|_2}{\|b\|_2}$	$\|\tilde{x}-x\|_\infty$
4	$5.0848e+006$	$8.1924e-017$	$1.2740e-011$	4.1657e-010	$7.7982e-012$
8	$1.4539e+017$	$3.6257e-015$	$1.4405e+002$	5.2712e+002	$1.0353e+002$
16	$1.1391e+019$	$9.0912e-015$	$4.9717e+002$	1.0355e+005	$3.8744e+002$
32	$2.8316e+018$	$2.3657e-014$	$1.1477e+003$	6.6987e+004	$1.5605e+003$
64	$1.0261e+019$	$1.0984e-014$	$6.3513e+002$	1.1270e+005	$9.7672e+002$
128	$3.3125e+019$	$5.8504e-014$	$6.0707e+003$	1.9379e+006	$1.0634e+004$

Table 5.2: Test problem 2: condition number, relative residual and relative error when different grid number N is used. The system $\mathbf{K}x = b$ is solved by Gaussian elimination with partial pivoting.

5.2.1 Discretization

Consider the 1-D boundary value problem (BVP)

$$\begin{cases} -(q(x)u'(x))' &= f(x), \quad x \in (0,1) \\ u(0) &= u(1) = 0 \end{cases} \tag{5.4}$$

where q and f are given functions and u is the solution we seek. It can be shown that the BVP (5.4) is equivalent to the variational formulation:

$$\begin{cases} \text{Find } u \text{ such that } u(0) = u(1) = 0 \text{ and} \\ \int_0^1 qu'v'\,dx = \int_0^1 fv\,dx, \quad \forall v \text{ such that } v(0) = v(1) = 0. \end{cases}$$

By applying finite element method at the collocation points $\{x_i = i/N,\ i = 0, 1, \ldots, N\}$ with "hat functions" $\{\phi_0, \phi_1, \ldots, \phi_N\}$ as our basis functions, our finite element problem is

$$\begin{cases} \text{Find } u_N \in V_h \text{ such that } u_N(0) = u_N(1) = 0 \text{ and} \\ \int_0^1 qu_N'v'\,dx = \int_0^1 fv\,dx, \quad \forall v \in V_h^0 \end{cases} \tag{5.5}$$

where

$$\begin{aligned} V_h &= \text{span}\,\{\phi_0, \phi_1, \ldots, \phi_N\}, \\ V_h^0 &= \text{span}\,\{\phi_1, \phi_2, \ldots, \phi_{N-1}\}. \end{aligned}$$

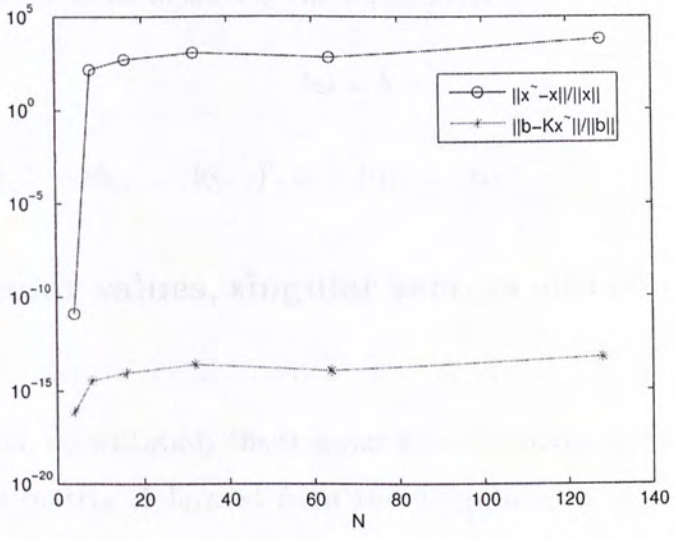

Figure 5.4: Test problem 2: relative error and relative residual when different grid number N is used. The system $\mathbf{K}x = b$ is solved by Gaussian elimination with partial pivoting.

Since $u_N \in V_h$, it can be written as

$$u_N(x) = \sum_{j=0}^{N} \alpha_j \phi_j(x).$$

To ensure that u_N satisfies the boundary conditions $u_N(0) = u_N(1) = 0$, we should have $\alpha_0 = \alpha_N = 0$. Thus $u_N(x)$ is of the form

$$u_N(x) = \sum_{j=1}^{N-1} \alpha_j \phi_j(x).$$

Since $\{\phi_i\}_{i=1}^{N-1}$ is the base of V_h^0, we have

$$\sum_{j=1}^{N-1} \alpha_j \int_0^1 q\phi_j'\phi_i'\,dx = \int_0^1 f\phi_i\,dx, \quad \text{for } i = 1, \ldots, N-1.$$

Let $a_{ij} = \int_0^1 q\phi_j'\phi_i'\,dx, b_i = \int_0^1 f\phi_i\,dx$, we have

$$\sum_{j=1}^{N-1} a_{ij}\alpha_j = b_i, \quad \text{for } i = 1, \ldots, N-1,$$

and so (5.5) is equivalent to solving the linear system

$$A\alpha = b$$

where $A = (a_{ij})$, $b = (b_1, \ldots, b_{N-1})'$, $\alpha = (\alpha_1, \ldots, \alpha_{N-1})'$.

5.2.2 Singular values, singular vectors and condition numbers

In this subsection, we will study the singular values, singular vectors and condition numbers of the matrix A formed from the discretization of a boundary value problem of the type (5.4), and compare them with those in the integral equation case in subsection 5.1.3. Set $N = 40$ and denote σ_i as the singular values of the matrix A. Figure 5.5 shows the distribution of the singular values and some of the right singular vectors for BVP (5.4) with q taken as $q(x) = 2 + \sin x$. We see that the singular values of A are much larger than those of $\mathbf{K}$ in figures 5.1, 5.2, and 5.3. The maximal singular value of A is very large and is greater than 10^4. We also observe that the smaller singular values correspond to the less oscillatory singular vectors, which is opposite to the integral equation case. Table 5.3 shows the condition numbers of A when different grid numbers N are used. We see that the condition number increases as the grid number increases and hence the more ill-conditioned the matrix A is. The ill-condition is due to the large maximal singular value. Compare with table 5.1, the condition number of A is smaller than that of $\mathbf{K}$ in test problems 1, 2 and 3, which means that the matrix A formed from the BVP (5.4) is less ill-conditioned than the matrix $\mathbf{K}$ formed from the first kind Fredholm integral equation (5.2).

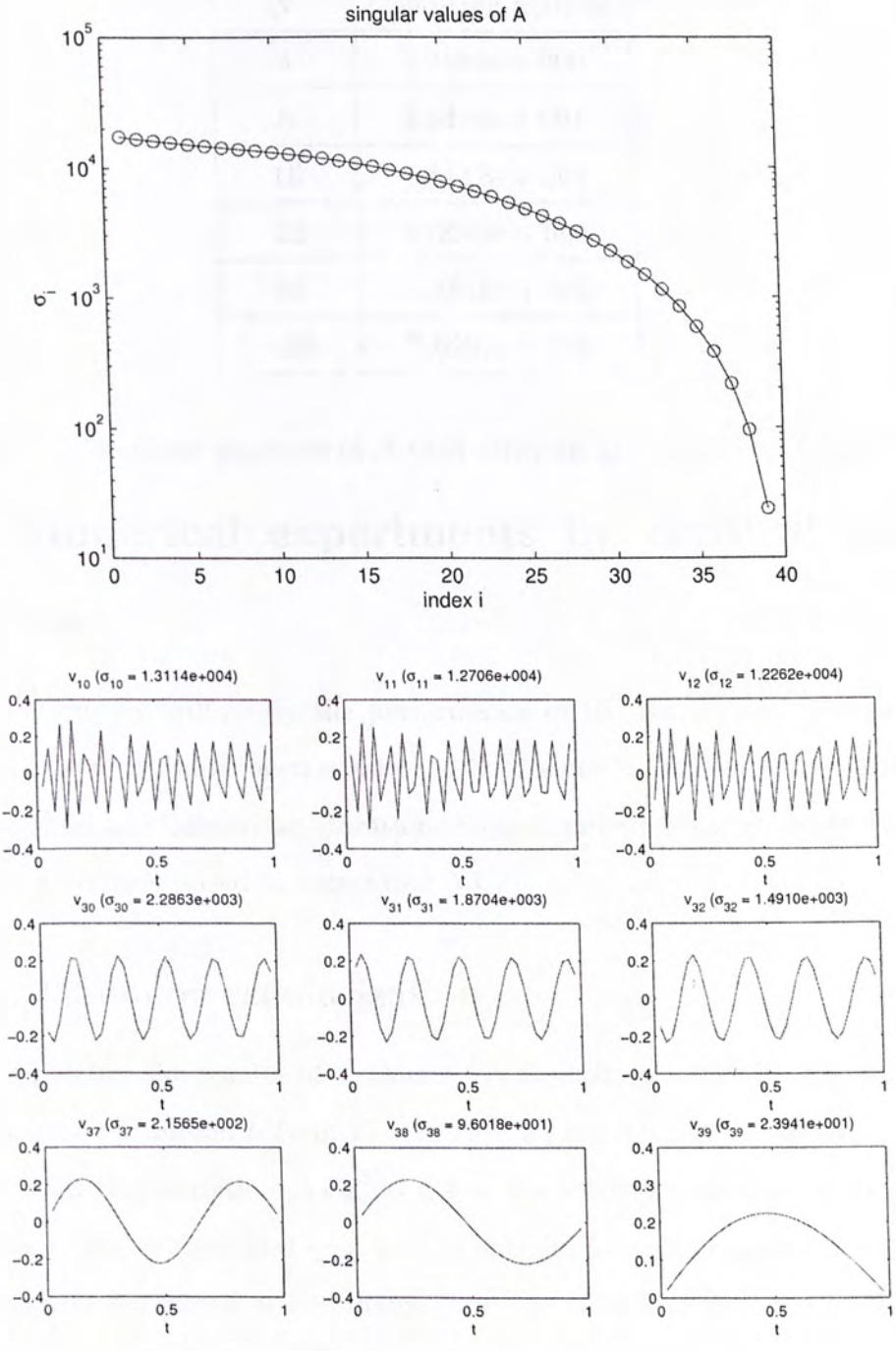

Figure 5.5: BVP: the top subplot shows the distribution of the singular values σ_i of the matrix A, and the subplots below it show the corresponding right singular vectors $\mathbf{v_i}$ for $i = 10, 11, 12, 30, 31, 32, 37, 38, 39$.

N	Condition number
4	$5.9564e+000$
8	$2.6402e+001$
16	$1.1142e+002$
32	$4.6088e+002$
64	$1.8816e+003$
128	$7.6202e+003$

Table 5.3: Condition numbers of A with different grid numbers N for BVP.

5.3 Numerical experiments by classical methods

In this section, we will study the performance of the three classical regularization methods which have been mentioned in chapter 3: Tikhonov regularization, TSVD method and Landweber iteration through numerical experiments using the three test problems posed in subsection 5.1.2.

5.3.1 Tikhonov regularization

Figure 5.6 shows the results of Tikhonov regularization with different choices of regularization parameter α from 10^{-7} to 10^{-2} for test problem 1. In this example, we have used the analytic data $\mathbf{g}^\delta = \mathbf{g} + \mathbf{e}$. We set the noise level to be $\delta = 1\%$ and let $N = 20$. We see that $\alpha = 10^{-4}$ is optimal for this example. Actually the regularization parameter α also depends on the noise level and the mesh size. If we change the noise level to $\delta = 6\%$, then from figure 5.7 we see that $\alpha = 10^{-4}$ will no longer be the optimal choice. Comparing figures 5.6 and 5.7, we can also see that the regularized solutions in figure 5.7 are poorer due to the higher noise level. Next we fix $N = 40$ and $\alpha = 10^{-6}$ to investigate how the noise level affect

the computed solution. Figure 5.8 shows the results of Tikhonov regularization of the same problem but with different choices of noise level δ from 10^{-5} to 10^{0}. We see that the relative error $\|u - u^{\delta}\|_2 / \|u\|_2$ increases nearly as a linear function of the noise level δ.

Consider test problem 2. In this example, we investigate the differences between using analytic data and synthetic data. We use Tikhonov regularization with $N = 20$ to demonstrate. In figure 5.9, we see that synthetic data generally lead to better approximate solutions, especially when the noise level δ is low. This is due to the fact that discretization errors are cancelled out when synthetic data are used. This is even more significant in cases that the data noise are small since discretization errors become relatively larger in such cases.

5.3.2 TSVD

Figure 5.10 shows the TSVD solutions for test problem 1. We have used the analytic data with $\delta = 1\%$ and have set $N = 20$. We see that the more singular values retained, the more oscillating the TSVD solution is, since the noise (which is usually of high-frequency) will be amplified by σ_i^{-1} in the regularized solution. It is found that the TSVD solution obtained by retaining the four largest SVD components is the best. Next consider test problem 3, in which the solution is discontinuous. We have used the exact data and have set $N = 20$. Figure 5.11 shows the TSVD solutions when different number of singular values are retained. We see that very few number of singular values retained yield overly smooth approximate solutions, and too many singular values retained make the approximate solutions highly oscillatory. It is found that the TSVD solution obtained by retaining the eight largest SVD components is the best.

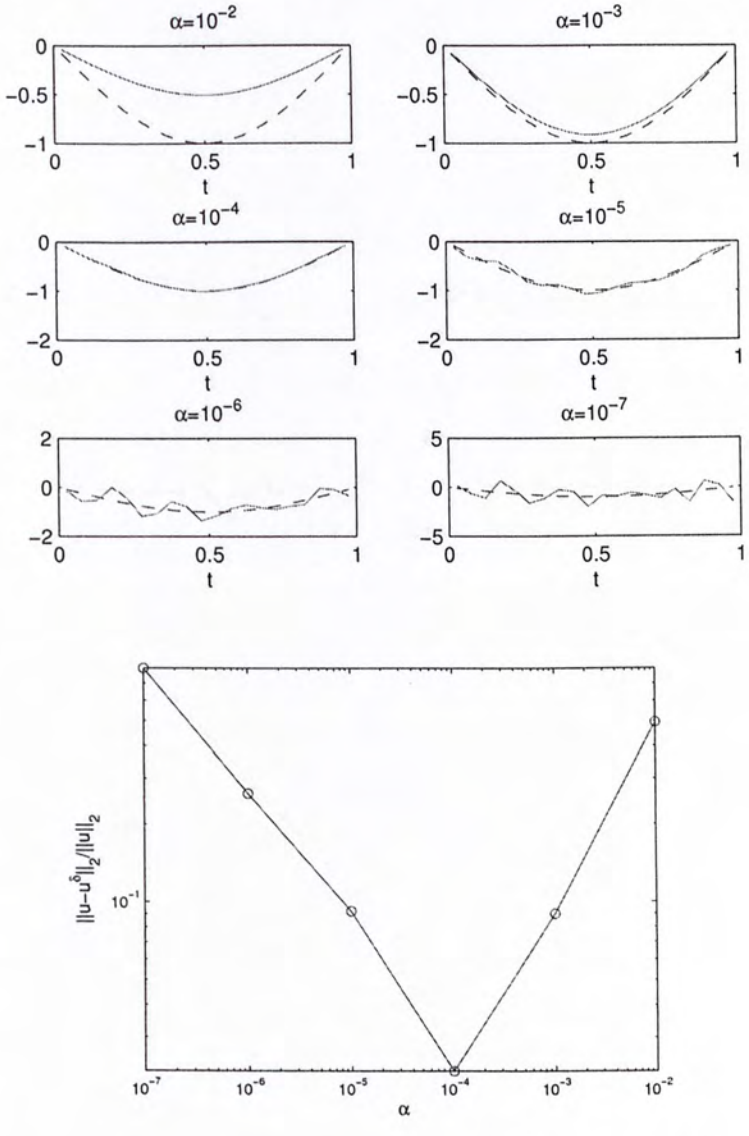

Figure 5.6: Test problem 1. The top 6 subplots show the Tikhonov regularized solutions and the exact solutions with different α, the solid line represents the regularized solution, while the dashed line represents the exact one. The bottom subplot shows the relative error between the regularized solution and exact solution with different choices of α. Here we have fixed $N = 20$ and $\delta = 1\%$.

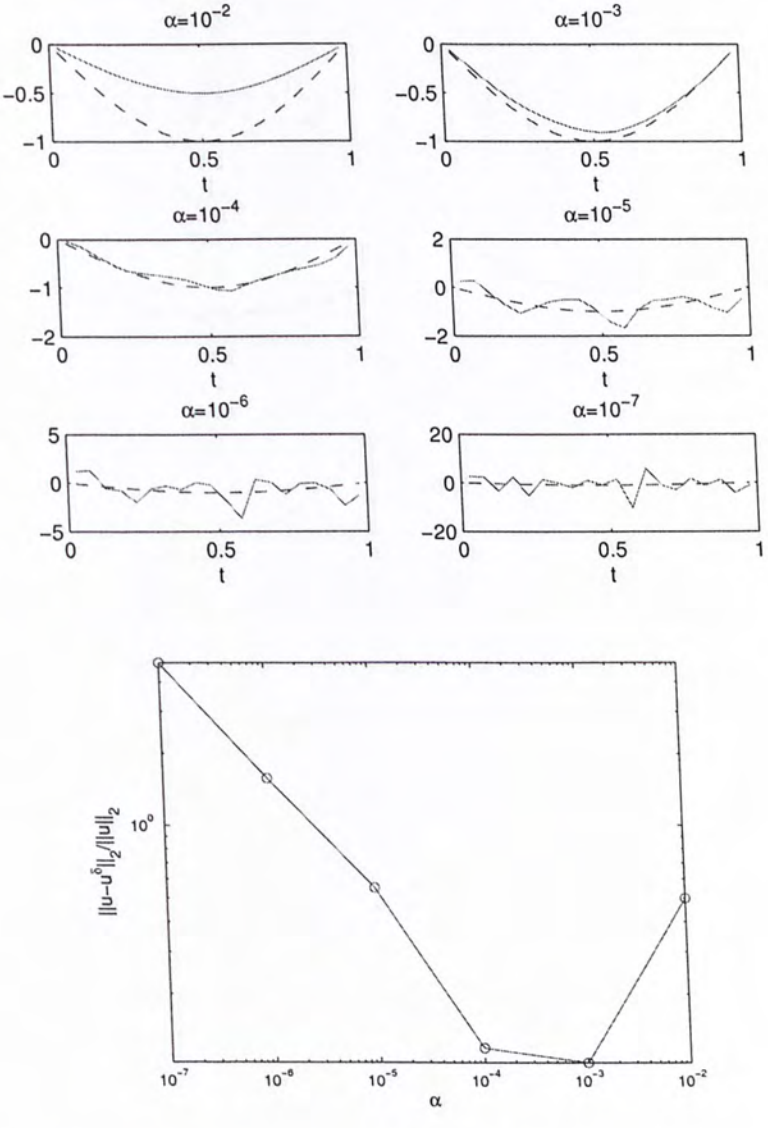

Figure 5.7: Test problem 1. The top 6 subplots show the Tikhonov regularized solutions and the exact solutions with different α, the solid line represents the regularized solution, while the dashed line represents the exact one. The bottom subplot shows the relative error between the regularized solution and exact solution with different choices of α. Here we have fixed $N = 20$ and $\delta = 6\%$.

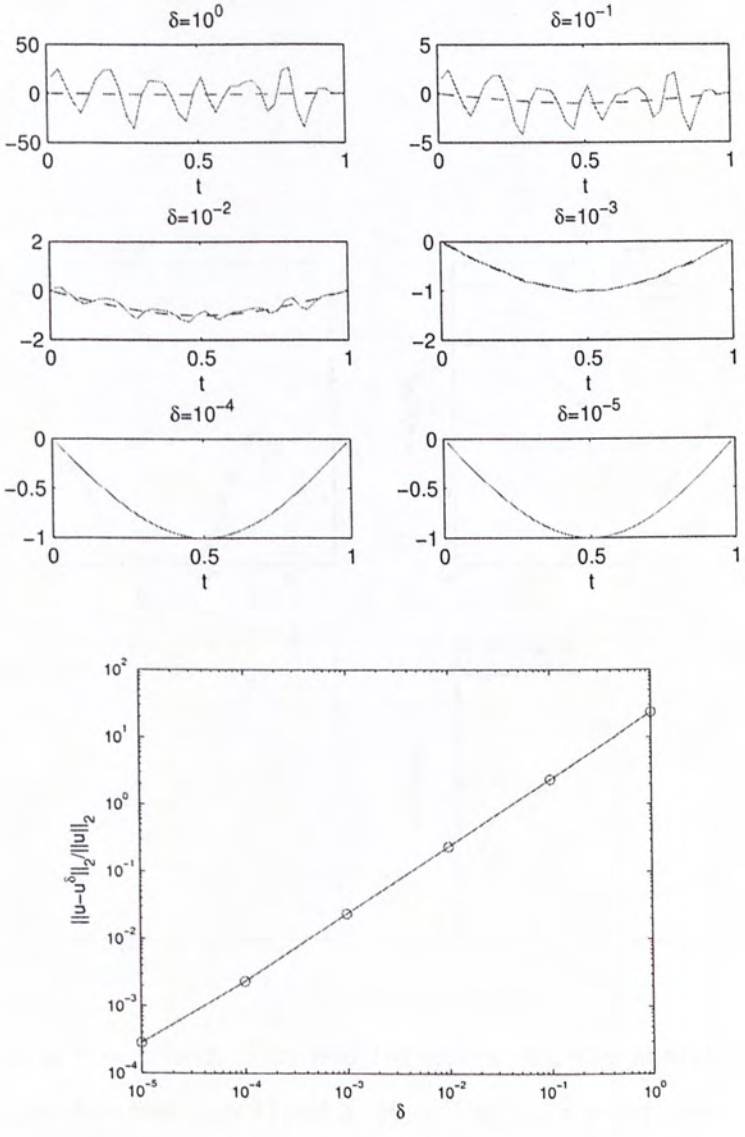

Figure 5.8: Test problem 1. The top 6 subplots show the Tikhonov regularized solutions and the exact solutions with different δ, the solid line represents the regularized solution, while the dashed line represents the exact one. The bottom subplot shows the relative error between the regularized solution and exact solution with different choices of δ. Here we have fixed $N = 40$ and $\alpha = 10^{-6}$.

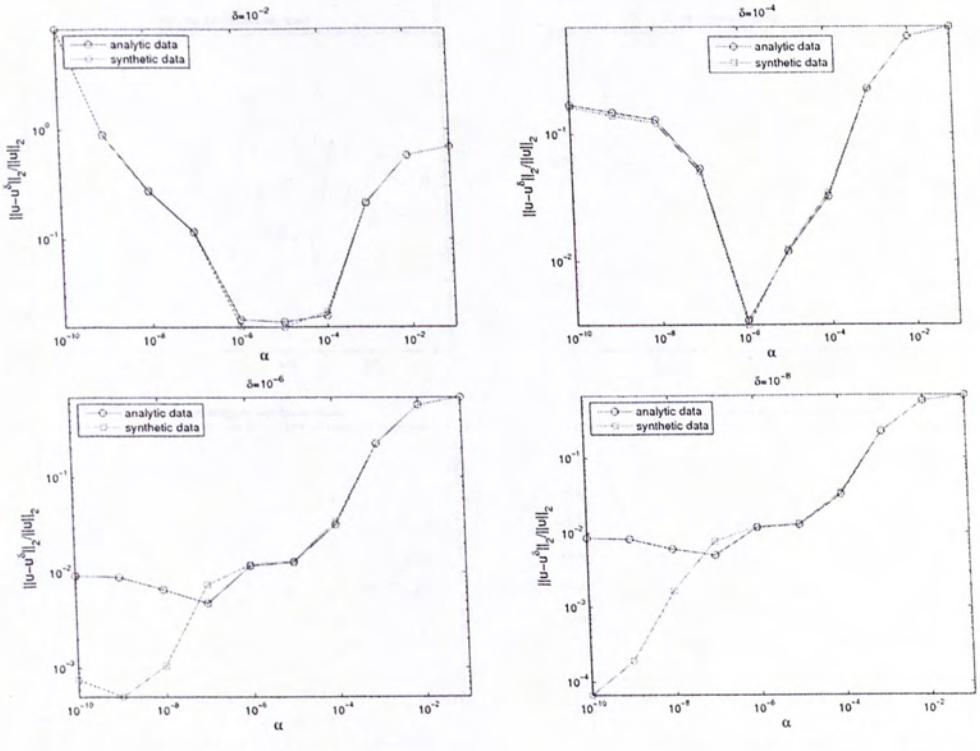

Figure 5.9: Test problem 2. The relative errors of using analytic data and synthetic data with different noise level δ. Here Tikhonov regularization is used with $N = 20$.

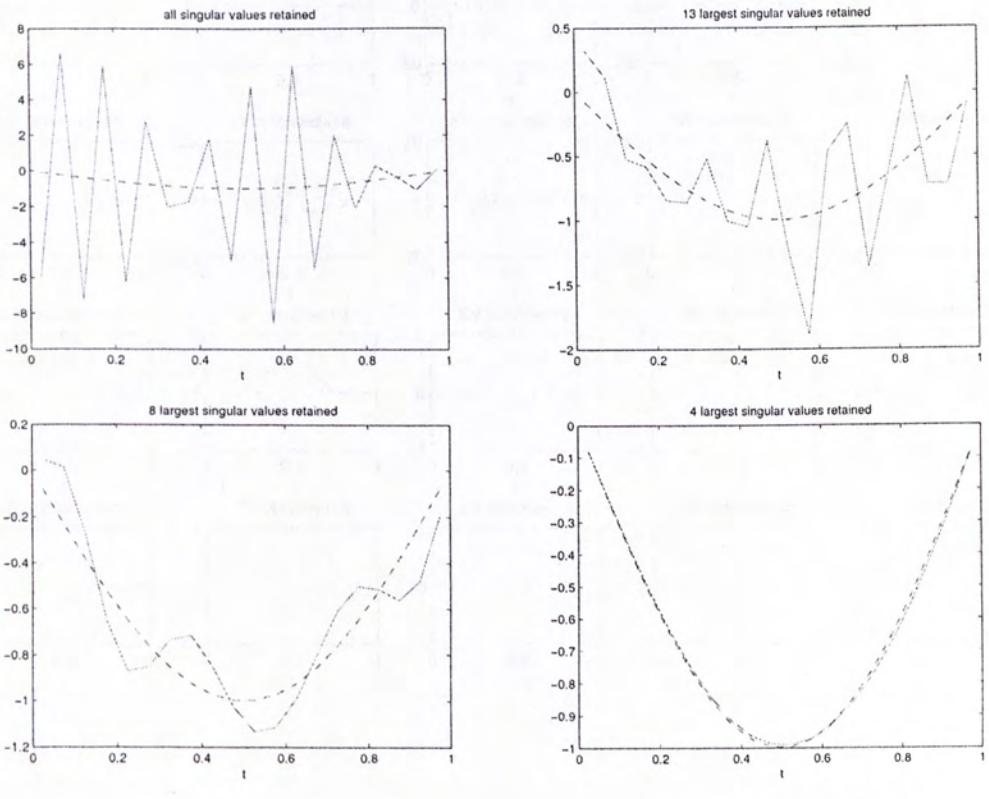

Figure 5.10: Test problem 1. The solid line represents the TSVD solution, while the dashed line represents the exact one. Here we have fixed $N = 20$ and $\delta = 1\%$.

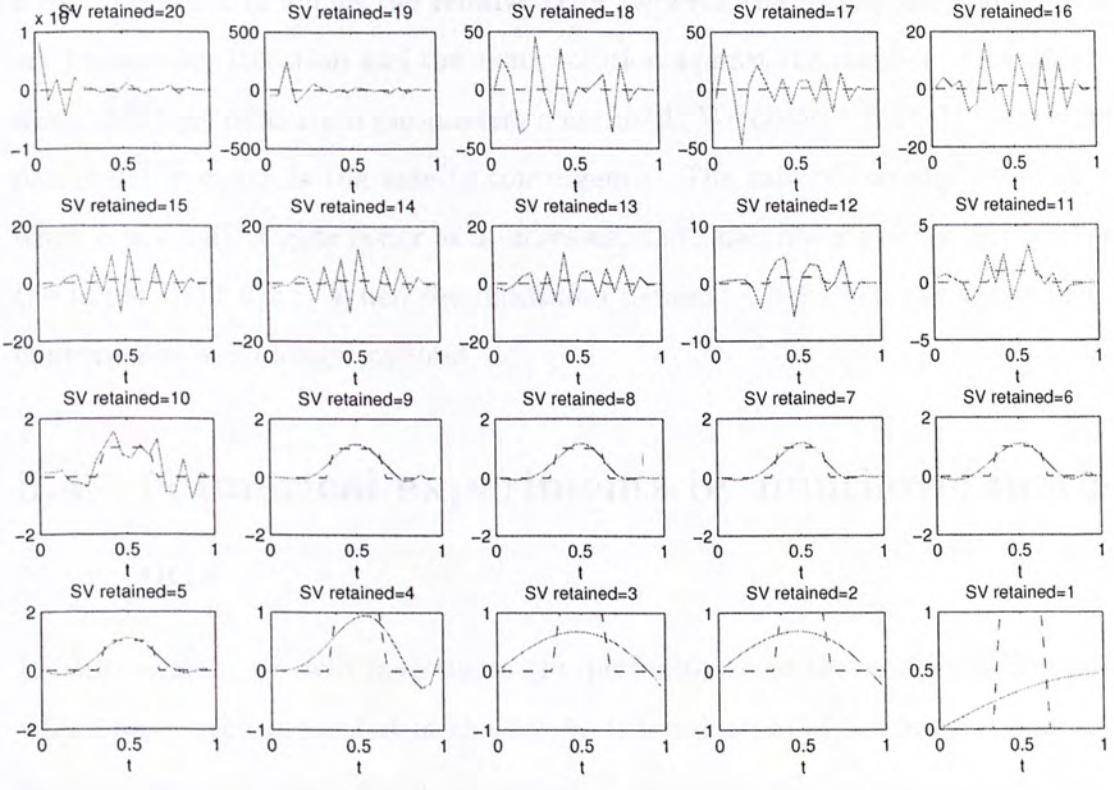

Figure 5.11: Test problem 3. The first subplot shows the TSVD solution when all the 20 singular values are retained, and the last subplot shows the TSVD solution when only the largest singular value is retained. The solid line represents the TSVD solution, while the dashed line represents the exact one. Here we have fixed $N = 20$ and $\delta = 0$.

5.3.3 Landweber iteration

Consider test problem 2. Again we use the analytic data with $\delta = 1\%$ and set $N = 20$. The largest singular value is found to be $\sigma_1 \approx 0.9447$ and so the relaxation parameter a in Landweber iteration (3.7) should satisfy $0 < a < 2/\sigma_1^2 \approx 2.241$. Figure 5.12 shows the relative error between the regularized solution using Landweber iteration and the exact solution against the number of iterations, when different relaxation parameters a are used. We observe that the relaxation parameter a controls the rate of convergence. The rate of convergence is slow when a is small, it gets faster as a increases, and slows down when it approaches the upper limit $2/\sigma_1^2$. When the relaxation parameter a exceeds the upper limit, convergence is no longer assured.

5.4 Numerical experiments by multilevel methods

In this section, we will investigate the performance of the multilevel methods which have been presented in chapter 4. Discretization of the integral operator and the general convergence of computed solutions will be introduced in subsection 5.4.1, in which the concrete subspaces and inner products are defined. Numerical results will be presented in subsection 5.4.2, and finally the effect of multilevel parameters on convergence will be studied in subsection 5.4.3.

5.4.1 General convergence

We follow [14] to introduce the discretization of the following integral operator $K : L_2[0,1] \to L_2[0,1]$ defined by

$$Ku(s) = \int_0^1 k(s,t)u(t)\,dt$$

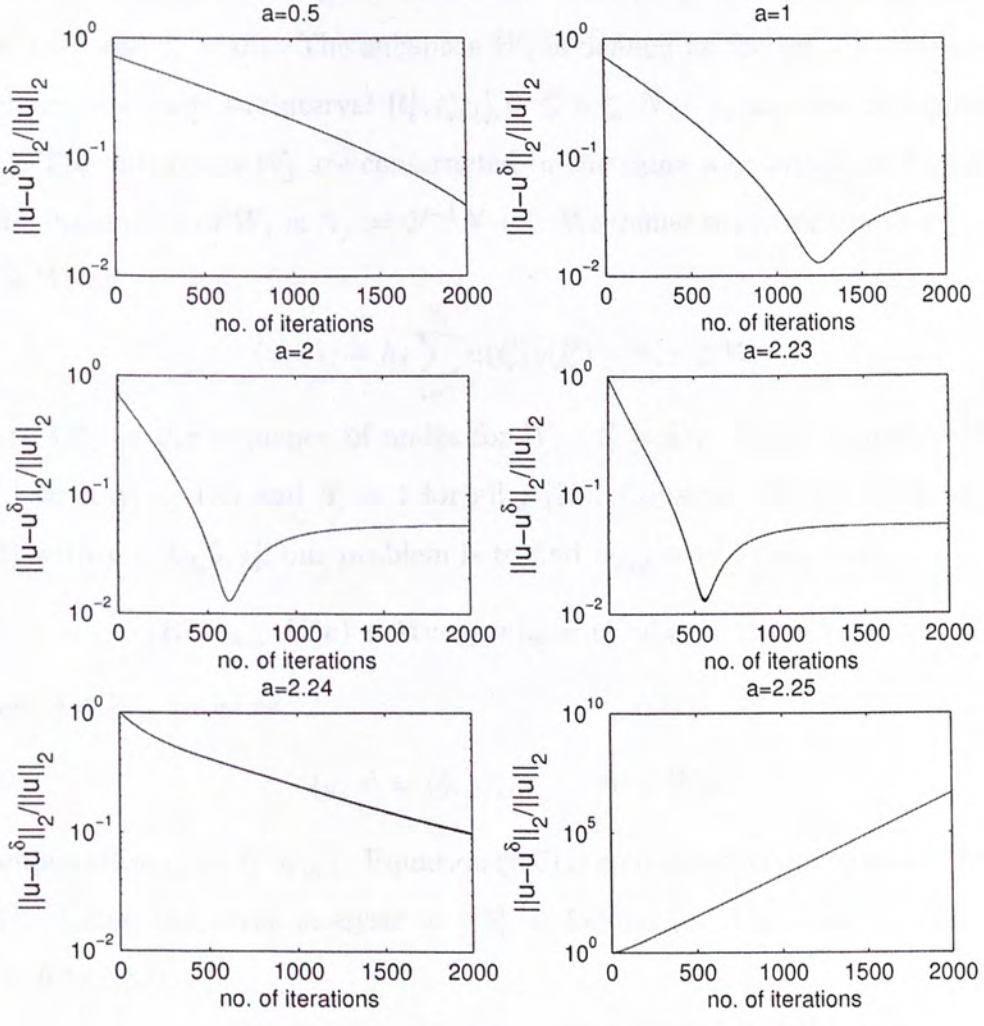

Figure 5.12: Test problem 2. The subplots show the relative errors versus the number of iterations when different relaxation parameters a are used in Landweber iteration. Here we have fixed $N = 20$ and $\delta = 1\%$.

where the kernel $k(\cdot,\cdot)$ is square integrable on $[0,1]\times[0,1]$ and $N(K^*)=\{0\}$. Thus $H_1 = H_2 = L_2[0,1] = \overline{R(K)}$ and the operator K is compact (please refer to section 4.1). We define the finite dimensional subspaces W_j consist of linear splines on a sequence of grids as follows. Let $N \geq 0$ be an integer and set $h_1 = 1/N$ and $t_i^1 = ih_1$. The subspace W_1 is defined as the set of functions that are linear on each subinterval $[t_i^1, t_{i+1}^1], 0 \leq i \leq N-1$, and are continuous on $[0,1]$. The subspaces W_j are constructed in the same way with $h_j = h_{j-1}/2$ and so the dimension of W_j is $N_j := 2^{j-1}N+1$. We define the inner product $\langle\cdot,\cdot\rangle_j$ on $W_j \times W_j$ as

$$\langle u, v\rangle_j = h_j \sum_{i=1}^{N_j} u(t_i^j)v(t_i^j) \qquad u, v \in W_j,$$

where $\{t_i^j\}$ is the sequence of nodes for $W_j : t_i^j = ih_j$. Then inequality (4.1) is valid with $\alpha_j = 1/6$ and $\beta_j = 1$ for all j [22]. Consider the first kind equation (5.2) with $g \in L_2[0,1]$, our problem is to find $w_{m,\lambda} \in W_m$ such that

$$(K^*w_{m,\lambda}, K^*v) + \lambda\langle w_{m,\lambda}, v\rangle_m = \langle \hat{g}, v\rangle_m, \qquad \forall v \in W_m, \tag{5.6}$$

where $\hat{g} \in W_m$ satisfies

$$\langle g, v\rangle = \langle \hat{g}, v\rangle_m, \qquad \forall v \in W_m.$$

Then we set $u_{m,\lambda} = K^*w_{m,\lambda}$. Equation (5.6) is equivalent to the operator equation (4.7). Using the error analysis in [15], it follows for $K^\dagger g = u^\dagger \in R(K^*)$, say $u^\dagger = K^*z$, that

$$\|u_{m,\lambda} - u^\dagger\| \leq \beta_m(\lambda^{1/2} + Ch_m^2)\|z\|$$

for some constant C. Thus for the choice $\lambda = \mathcal{O}(h_m^4)$, we have the error estimate

$$\|u_{m,\lambda} - u^\dagger\| \leq Ch_m^2\|z\|.$$

5.4.2 Numerical results

In this subsection, we will give some numerical results of the multilevel methods which have been presented in section 4.3. The total number of levels m is chosen

to be $m = 4$ and the four level operator C_4 is applied to test problem 1 in our numerical experiments. We have used the analytic data and have set $N = 2$. The multilevel parameters are optimal and are found by *trial and error* in this subsection. The choices of multilevel parameters will be discussed in the next subsection. Figures 5.13 and 5.14 show the results of Landweber iteration with nonzero regularization parameter λ (please refer to subsection 4.3.1). The parameters are chosen as $\lambda = 10h_4^4, \xi_j = (\lambda + h_j^4)^{-1}$, and $\lambda = 35h_4^4, \xi_j = (\lambda + h_j^4)^{-1}$ respectively. The noise levels are set to be 1% and 10% respectively. We observe, in both figures, that the relative error attains its minimum at around 5 iterations and is almost kept constant thereafter. Therefore we do not need to consider much about the number of iterations and we may just simply take, say, 100 iterations. The approximate solution in figure 5.14 is poorer due to the higher noise level. The results of Landweber iteration with zero regularization parameter λ (please refer to subsection 4.3.2) are shown in figures 5.15 and 5.16. The parameters are chosen as $\lambda = 10^3 h_4^4$ and $\xi_j = (\lambda + h_j^4)^{-1}$, and the number of iterations are 10 and 200 respectively. From figure 5.15, we see that the approximate solution using Landweber iteration with zero regularization parameter is better when the noise level is 10% compared to the case in figure 5.14 and the relative error attains its minimum at round 8 iterations. However, comparing figures 5.15 and 5.16, we see that the relative error grows up if we iterate too many times, and so we need to consider about the number of iterations when using the method Landweber iteration with zero regularization parameter.

Figures 5.17 and 5.18 show the results of CG method with nonzero regularization parameter λ. The parameters are chosen as $\lambda = 10h_4^4, \xi_j = (\lambda + h_j^4)^{-1}$, and $\lambda = 45h_4^4, \xi_j = (\lambda + h_j^4)^{-1}$ respectively. The noise levels are set to be 1% and 10% respectively. We observe that the relative error oscillates at the beginning and gets steady as the number of iterations increases, and the approximation in figure 5.18 is worse as the noise level is higher. Again we do not need to consider much

about the number of iterations and we may simply take 100 iterations. The results of CG method with zero regularization parameter λ is shown in figure 5.19. The parameters are chosen as $\lambda = 10^6 h_4^4$ and $\xi_j = (\lambda + h_j^4)^{-1}$. We see that the approximation is better with zero regularization parameter when the noise level is 10% compared to the results in figure 5.18. However the approximate solution becomes worse and more oscillatory when the number of iterations increases. We see that one iteration is the best.

Figure 5.20 shows the results of full multilevel algorithm (please refer to subsection 4.3.3) with different choices of noise level δ and parameter λ. The number of iterations per level is chosen such that the relative difference between successive iterates is no greater than some TOL. That is, we terminate on level j when $\|u_j^m - u_j^{m-1}\|_2 \leq TOL\|u_j^m\|_2$. In this example, we have set $TOL = 10^{-3}$. The maximum number of iterations permitted is set to be 20 on each level. The parameter ξ_j is chosen as $\xi_j = (10^4\lambda + h_j^4)^{-1}$ in each case. We observe that the approximation becomes worse as the noise level increases and we need to use different parameters λ when the noise levels are different.

To conclude, the multilevel parameters depend on the method we used, and also the noise level. We will see how the parameters affect the convergence in the next subsection. The results of the three methods: Landweber iteration with zero regularization parameter, CG method with zero regularization parameter and full multilevel algorithm are similar and are better than those of the other two methods: Landweber iteration with nonzero regularization parameter and CG method with nonzero regularization parameter, when the noise level is higher. We do not need to consider much about the number of iterations except when using the method Landweber iteration with zero regularization parameter.

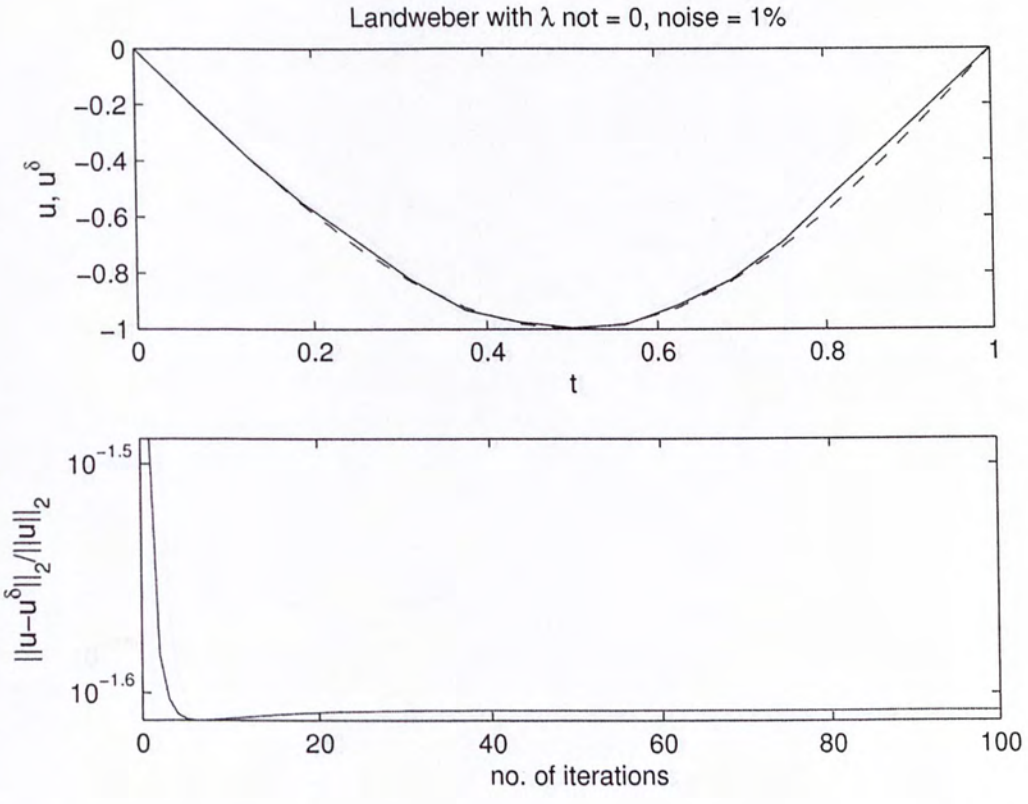

Figure 5.13: Nonzero regularization parameter - Landweber iteration. Top: the solid line represents the approximate solution with C_4 as the preconditioner (100 iterations), while the dashed line represents the exact solution. Bottom: the relative error against the number of iterations. Here $\lambda = 10h_4^4$, $\xi_j = (\lambda + h_j^4)^{-1}$ and $\delta = 1\%$.

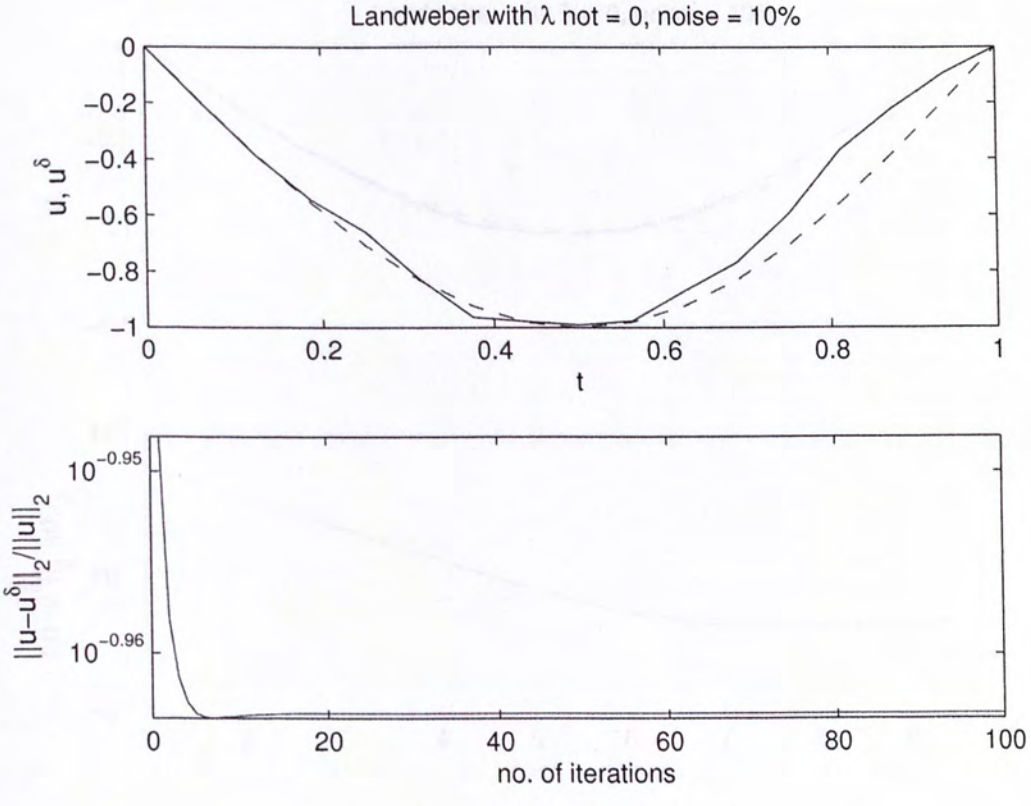

Figure 5.14: Nonzero regularization parameter - Landweber iteration. Top: the solid line represents the approximate solution with C_4 as the preconditioner (100 iterations), while the dashed line represents the exact solution. Bottom: the relative error against the number of iterations. Here $\lambda = 35h_4^4$, $\xi_j = (\lambda + h_j^4)^{-1}$ and $\delta = 10\%$.

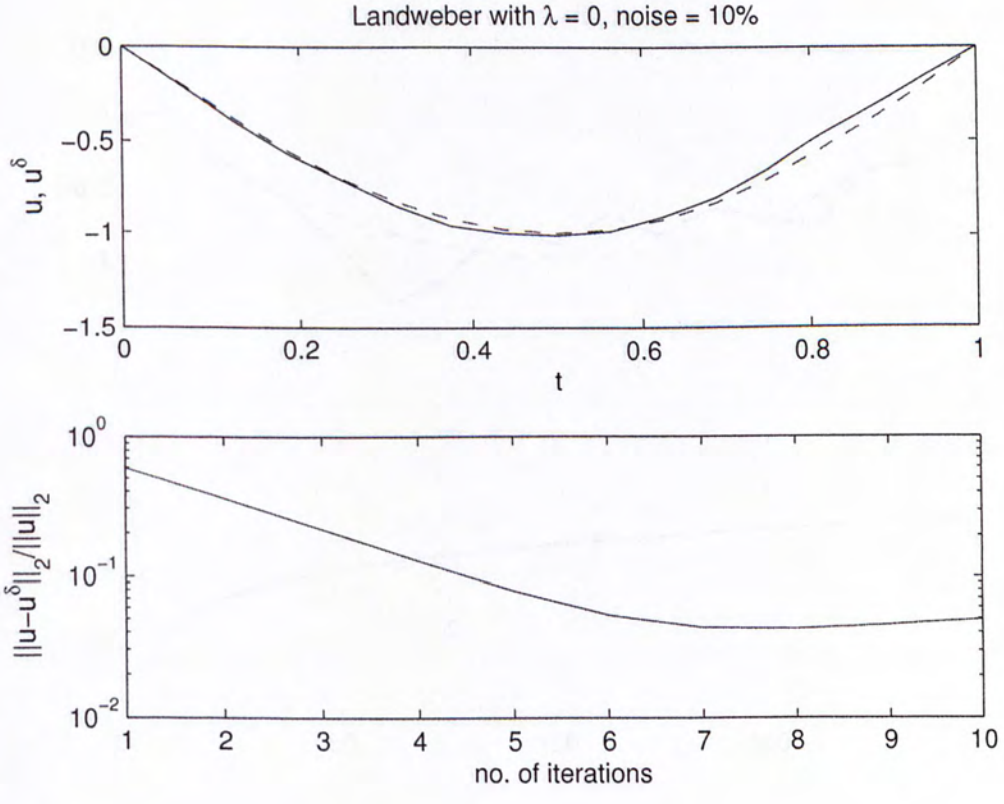

Figure 5.15: Zero regularization parameter - Landweber iteration. Top: the solid line represents the approximate solution with C_4 as the preconditioner (10 iterations), while the dashed line represents the exact solution. Bottom: the relative error against the number of iterations. Here $\lambda = 10^3 h_4^4$, $\xi_j = (\lambda + h_j^4)^{-1}$ and $\delta = 10\%$.

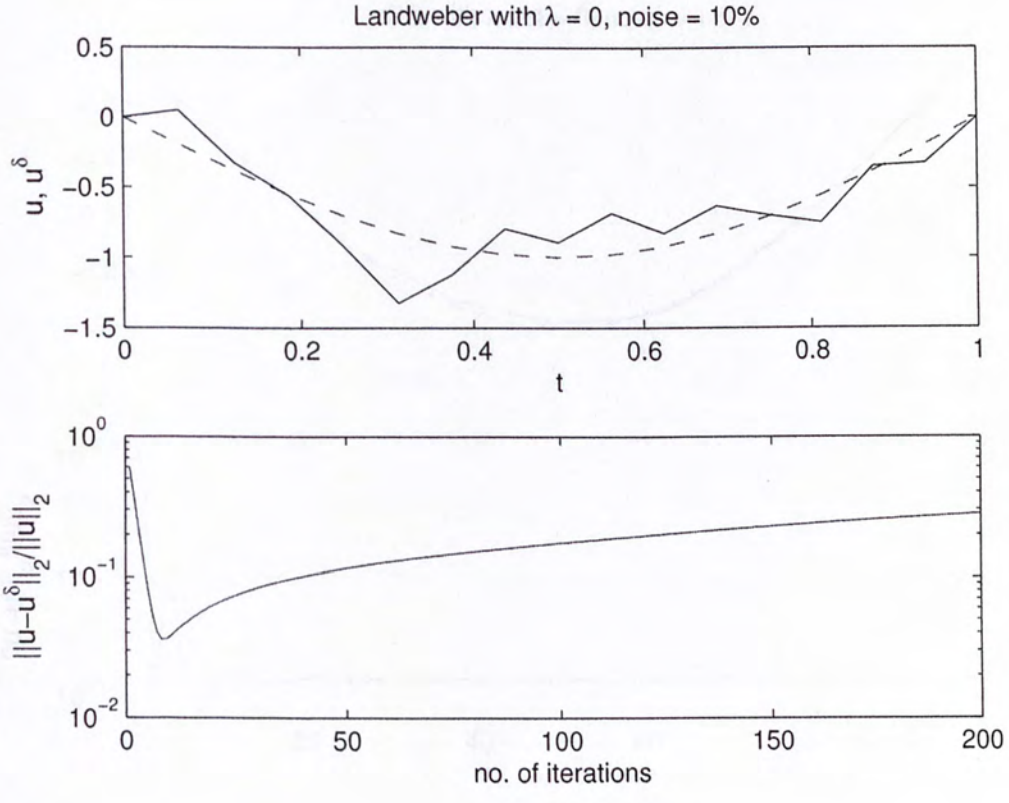

Figure 5.16: Zero regularization parameter - Landweber iteration. Top: the solid line represents the approximate solution with C_4 as the preconditioner (200 iterations), while the dashed line represents the exact solution. Bottom: the relative error against the number of iterations. Here $\lambda = 10^3 h_4^4$, $\xi_j = (\lambda + h_j^4)^{-1}$ and $\delta = 10\%$.

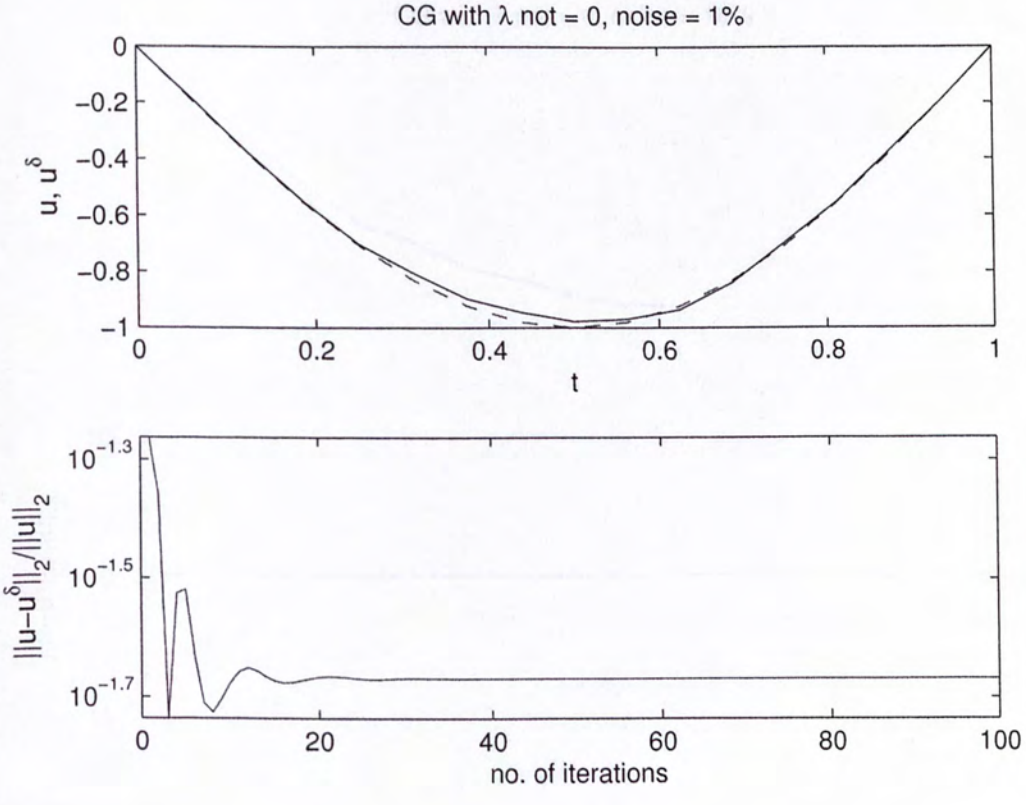

Figure 5.17: Nonzero regularization parameter - CG method. Top: the solid line represents the approximate solution with C_4 as the preconditioner (100 iterations), while the dashed line represents the exact solution. Bottom: the relative error against the number of iterations. Here $\lambda = 10h_4^4$, $\xi_j = (\lambda + h_j^4)^{-1}$ and $\delta = 1\%$.

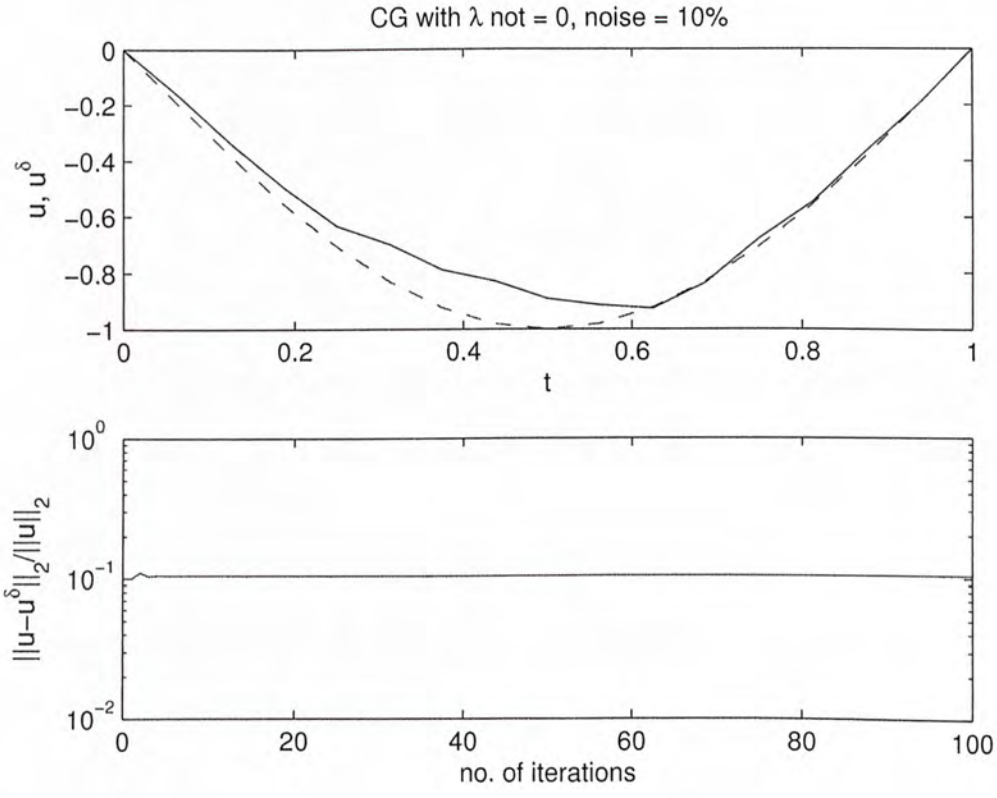

Figure 5.18: Nonzero regularization parameter - CG method. Top: the solid line represents the approximate solution with C_4 as the preconditioner (100 iterations), while the dashed line represents the exact solution. Bottom: the relative error against the number of iterations. Here $\lambda = 45h_4^4$, $\xi_j = (\lambda + h_j^4)^{-1}$ and $\delta = 10\%$.

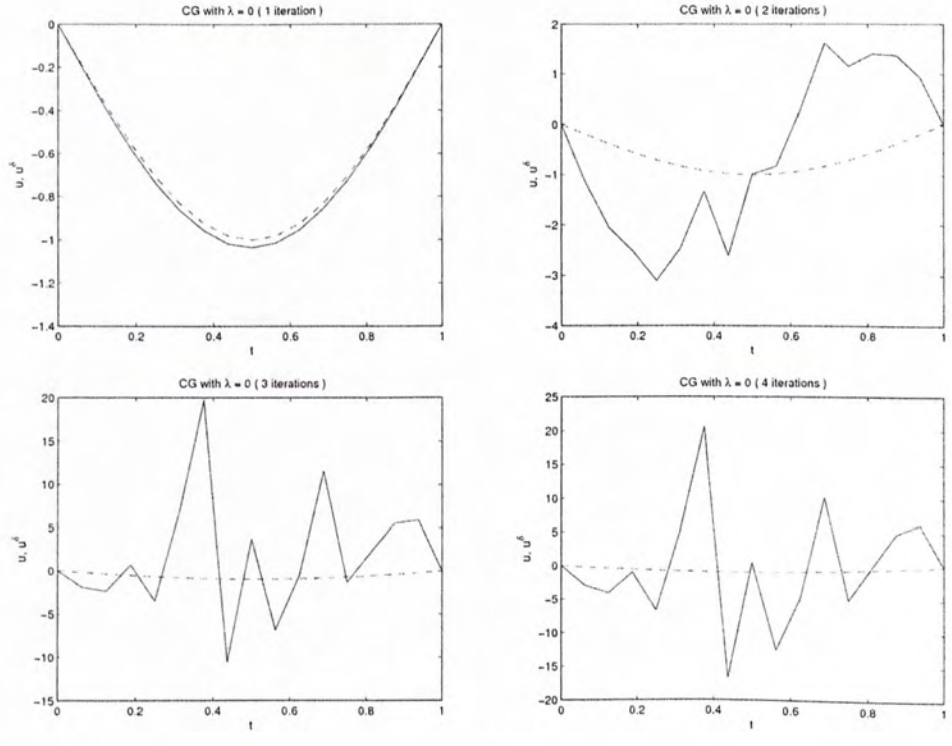

Figure 5.19: Zero regularization parameter - CG method with different numbers of iterations. The solid line represents the approximate solution with C_4 as the preconditioner, while the dashed line represents the exact solution. Here $\lambda = 10^6 h_4^4$, $\xi_j = (\lambda + h_j^4)^{-1}$ and $\delta = 10\%$.

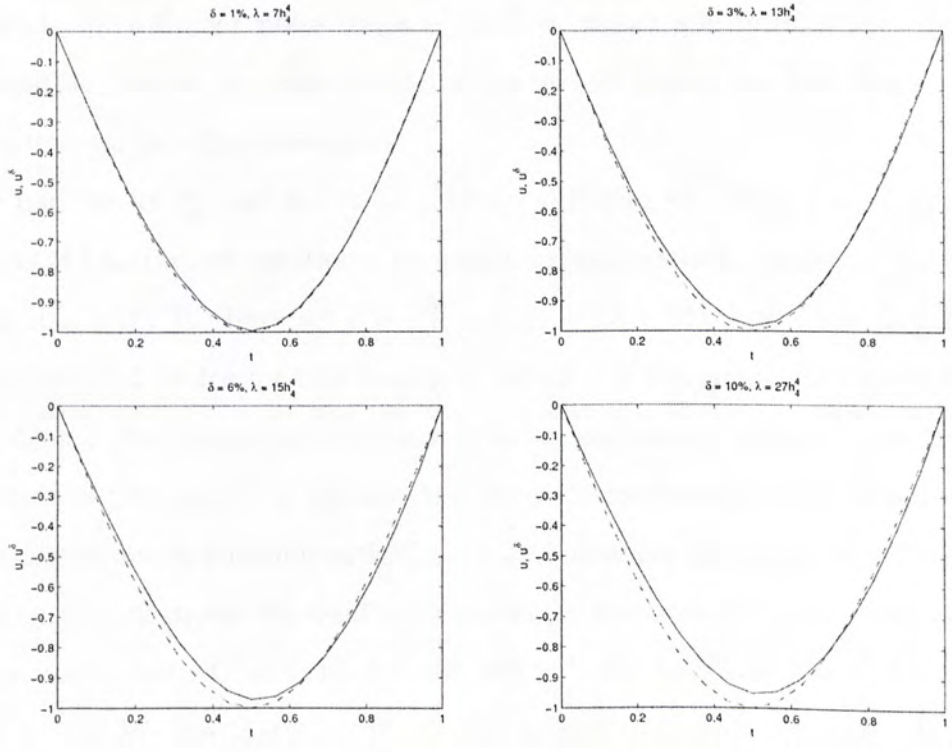

Figure 5.20: Full multilevel algorithm with different noise levels δ and parameters λ: the solid line represents the approximate solution with C_4 as the preconditioner, while the dashed line represents the exact solution. Here $\xi_j = (10^4\lambda + h_j^4)^{-1}$.

5.4.3 Effect of multilevel parameters on convergence

In subsection 5.4.1, we see that the choice $\lambda = \mathcal{O}(h_m^4)$ leads to the error estimate

$$\|u_{m,\lambda} - u^\dagger\| \leq Ch_m^2\|z\|.$$

So in subsection 5.4.2, we choose $\lambda = Ch_m^4$ for some C according to different methods, or different noise levels. The C's chosen are optimal and are found by trial and error. In this subsection, we would like to see how the multilevel parameters affect the convergence.

First we fix ξ_j and try $\lambda = Ch_4^4$ for different C. Figure 5.21 shows the results of Landweber iteration with nonzero regularization parameter λ with $C = 1, 10, 100, 1000$. We have set $\delta = 1\%$ and $\xi_j = (\lambda + h_j^4)^{-1}$. We see that $C = 10$ is the optimal choice for this example. When λ is too small, for example in the case $C = 1$, the computed solution will be too oscillatory; when λ is too large, for example in the cases $C = 100$ or 1000, the computed solution will be too smooth. This phenomenon coincide with that in Tikhonov regularization (see figure 5.6).

Figure 5.22 shows the results of Landweber iteration with zero regularization parameter λ with $C = 1, 10, 10^2, 10^3, 10^4, 10^5$. We have set $\delta = 10\%$ and $\xi_j = (\lambda + h_j^4)^{-1}$. We see that $C = 10^3$ is the optimal choice for this example. In the previous example, we see that for too large or too small λ, even the best computed solution is still unsatisfactory. In this example, for large λ, the best computed solution is much better but requires more iterations to attain the same level of accuracy (the last three subplots). However, it has been found that the relative error grows up if we iterate too many times (as in figure 5.16) and so we need to find out the optimal number of iterations. This is definitely a tradeoff!

Figure 5.23 shows the results of CG method with nonzero regularization parameter λ with $C = 1, 10, 100, 1000$. We have set $\delta = 1\%$ and $\xi_j = (\lambda + h_j^4)^{-1}$. The results is similar to the Landweber case with nonzero regularization parameter λ (figure 5.21) except that the relative error oscillates in the case $C = 1$.

Figure 5.24 shows the results of CG method with zero regularization parameter λ with $C = 10^1, 10^2, 10^3, 10^4, 10^5, 10^6$. We have set $\delta = 10\%$ and $\xi_j = (\lambda + h_j^4)^{-1}$. It has been found that the computed solution will become worse and more oscillatory when the number of iterations increases (as in figure 5.19) and one iteration is the optimal. We see that the case $C = 10^3$ is the best, the cases $C = 10$ and $C = 10^2$ are not satisfactory, but for other C greater than 10^3 the computed solution does not change much and is acceptable. We conclude that we may just simply take C large enough, say $C \geq 10^3$, for this example.

Figure 5.25 shows the results of full multilevel algorithm with C=0.01, 0.1, 1, 10, 100, 1000. We have set $\delta = 3\%$ and $\xi_j = (10^4\lambda + h_j^4)^{-1}$. We see that the results is again similar to the Landweber case with nonzero regularization parameter λ (figure 5.21) and the Tikhonov regularization phenomenon happens.

Next, we fix λ and change ξ_j to see how the parameter ξ_j affects the convergence. We try $\xi_j = (\tilde{C}\lambda + h_j^4)^{-1}$ for different $\tilde{C}$ and $\xi_j = \lambda^{-1}$ (which violates (4.2)). Figure 5.26 shows the results of Landweber iteration with nonzero regularization parameter λ with $\tilde{C} = 0.01, 0.1, 1, 10, 100$ and $\xi_j = \lambda^{-1}$. We have set $\delta = 1\%$ and $\lambda = 10h_4^4$. We see that too small the ξ_j will make the computed solution too oscillatory (the first two subplots in which $\tilde{C} = 0.01, 0.1$) but large ξ_j will not make big difference to the computed solution (the next three subplots in which $\tilde{C} = 1, 10, 100$). Therefore we may just simply take $\tilde{C} \geq 1$. From the last subplot, we see that even (4.2) is violated, the computed solution is still acceptable. Similar results is obtained for Landweber iteration with zero regularization parameter λ in figure 5.27, in which we have set $\delta = 10\%$ and $\lambda = 10^3 h_4^4$.

Figure 5.28 shows the results of CG method with nonzero regularization parameter λ with $\tilde{C} = 0.01, 0.1, 1, 10, 100$ and $\xi_j = \lambda^{-1}$. We have set $\delta = 1\%$ and $\lambda = 10h_4^4$. We see that for too small ξ_j ($\tilde{C} = 0.01, 0.1$), the relative error oscillates very much and so the computed solution is not reliable. For larger ξ_j ($\tilde{C} = 1, 10, 100$), the relative error oscillates at the beginning and gets steady as

the number of iterations increases. The relative error becomes almost constant after about 20 iterations. Therefore we may simply take $\tilde{C} \geq 1$ and iterates more than 20 times.

Figure 5.29 shows the results of CG method with zero regularization parameter λ with $\tilde{C} = 0.01, 0.1, 1, 10, 100$ and $\xi_j = \lambda^{-1}$. We have set $\delta = 10\%$ and $\lambda = 10^3 h_4^4$. Again we see that too small the ξ_j ($\tilde{C} = 0.01, 0.1$) will make the computed solution too oscillatory but large ξ_j ($\tilde{C} = 1, 10, 100$) will not make big difference to the computed solution. Therefore we may just simply take $\tilde{C} \geq 1$. From the last subplot, we see that the computed solution is still good even (4.2) is violated. Similar results is obtained for full multilevel algorithm in figure 5.30, in which we have set $\delta = 10\%$ and $\lambda = 27 h_4^4$.

In conclusion, CG method with zero regularization parameter is superior among these methods: the computed solution is good even for large noise level ($\delta = 10\%$), we do not need to consider much about the number of iterations (one iteration is the optimal), and the multilevel parameters λ and ξ_j are easy to choose (we may simply choose large C and $\tilde{C}$, say $C \geq 10^3$ and $\tilde{C} \geq 1$).

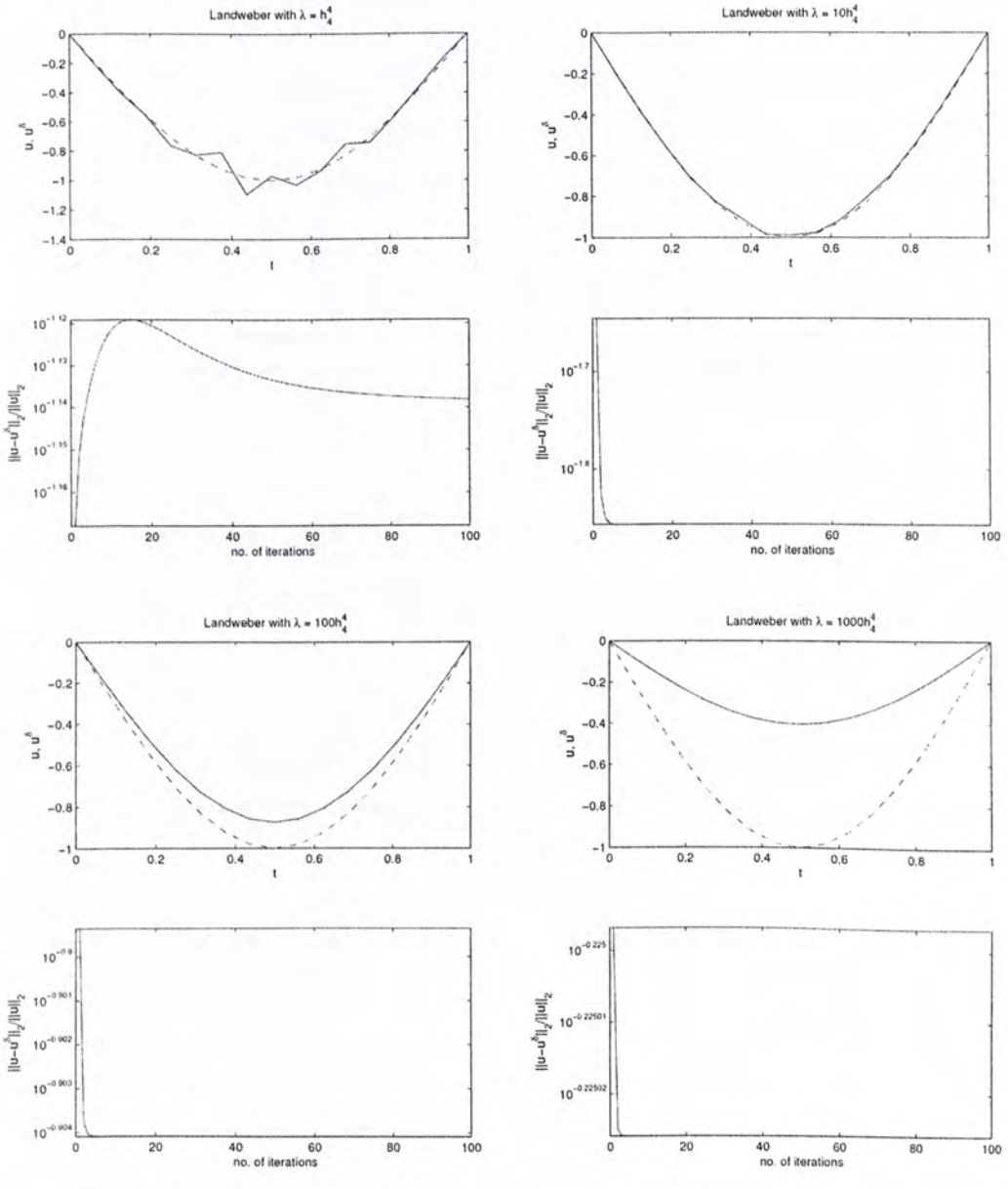

Figure 5.21: Nonzero regularization parameter - Landweber iteration with different λ, and relative error. The solid line represents the approximate solution with C_4 as the preconditioner (100 iterations), while the dashed line represents the exact solution. Here $\delta = 1\%$ and $\xi_j = (\lambda + h_j^4)^{-1}$.

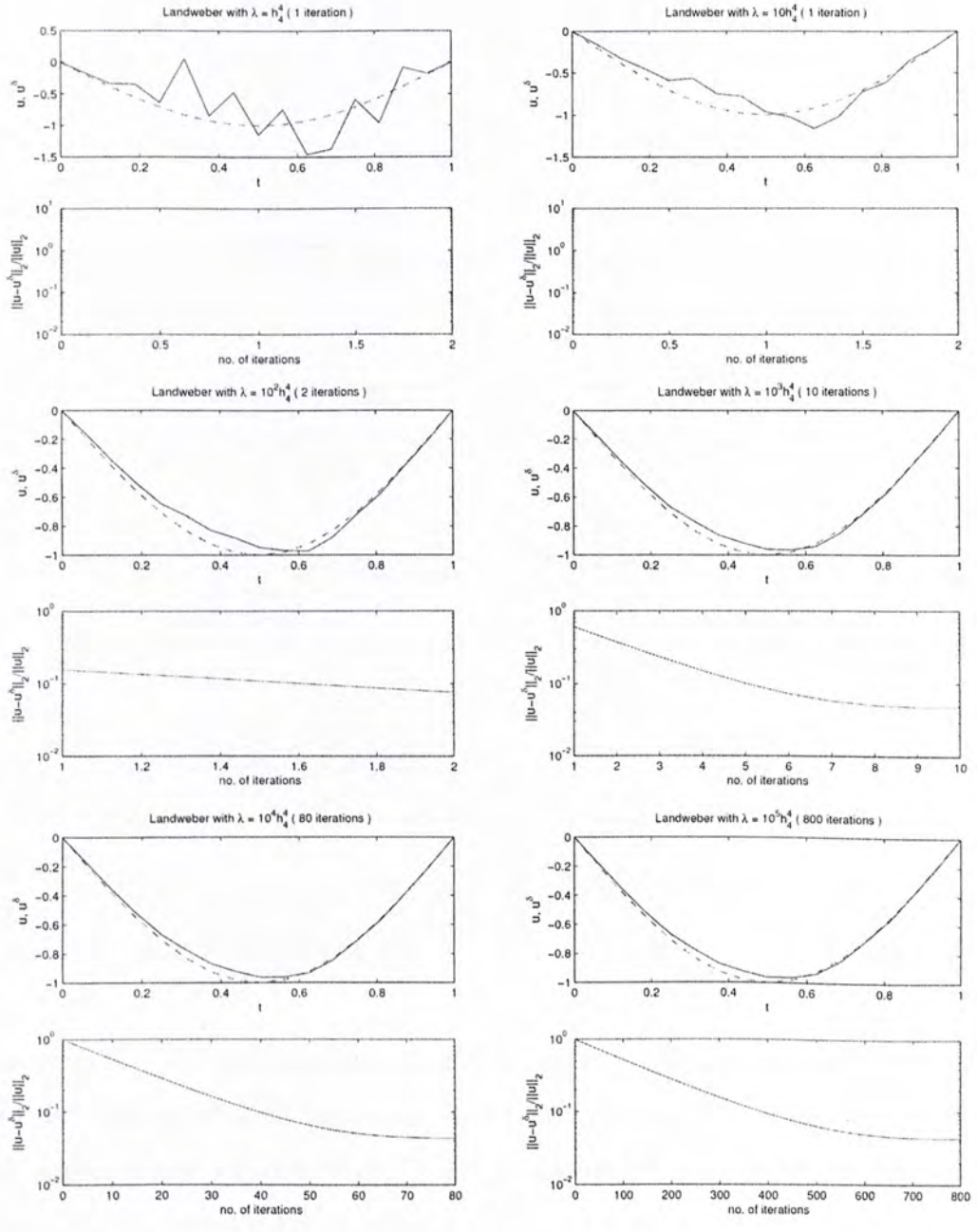

Figure 5.22: Zero regularization parameter - Landweber iteration with different λ and optimal number of iterations, and relative error. The solid line represents the approximate solution with C_4 as the preconditioner, while the dashed line represents the exact solution. Here $\delta = 10\%$ and $\xi_j = (\lambda + h_j^4)^{-1}$.

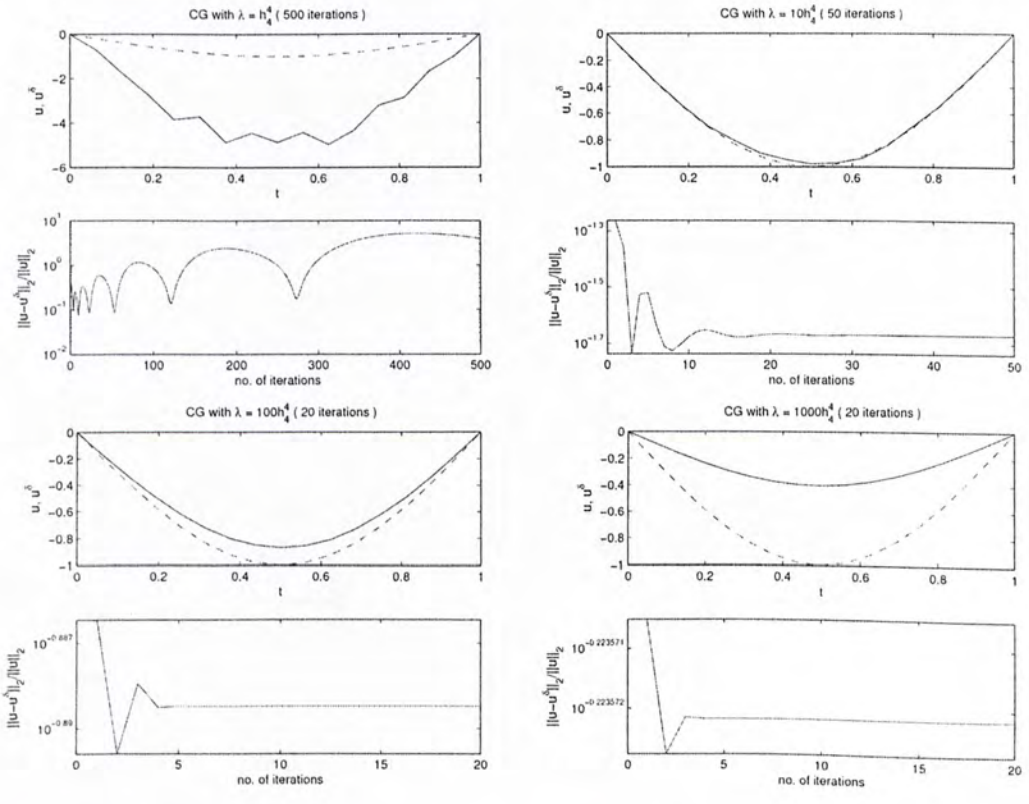

Figure 5.23: Nonzero regularization parameter - CG method with different λ and different number of iterations, and relative error. The solid line represents the approximate solution with C_4 as the preconditioner, while the dashed line represents the exact solution. Here $\delta = 1\%$ and $\xi_j = (\lambda + h_j^4)^{-1}$.

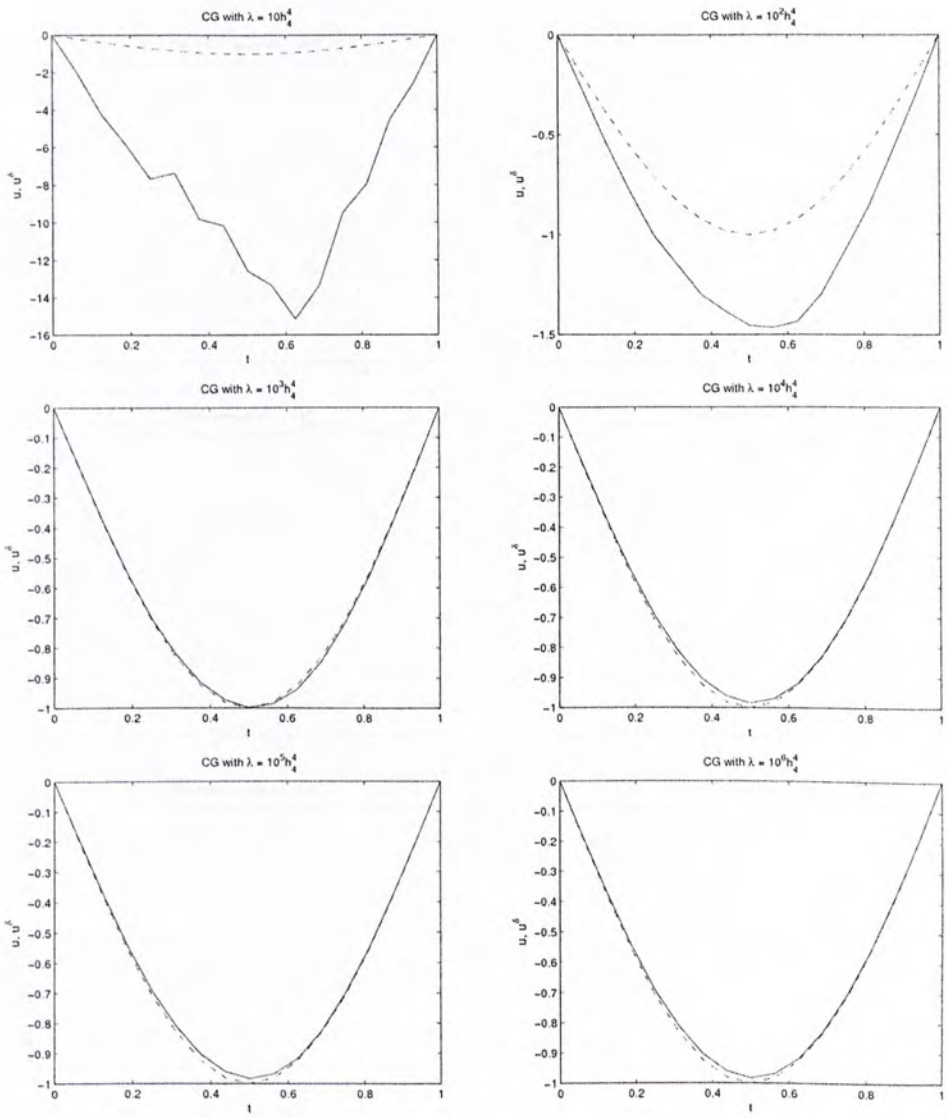

Figure 5.24: Zero regularization parameter - CG method with different λ. The solid line represents the approximate solution with C_4 as the preconditioner (1 iteration), while the dashed line represents the exact solution. Here $\delta = 10\%$ and $\xi_j = (\lambda + h_j^4)^{-1}$.

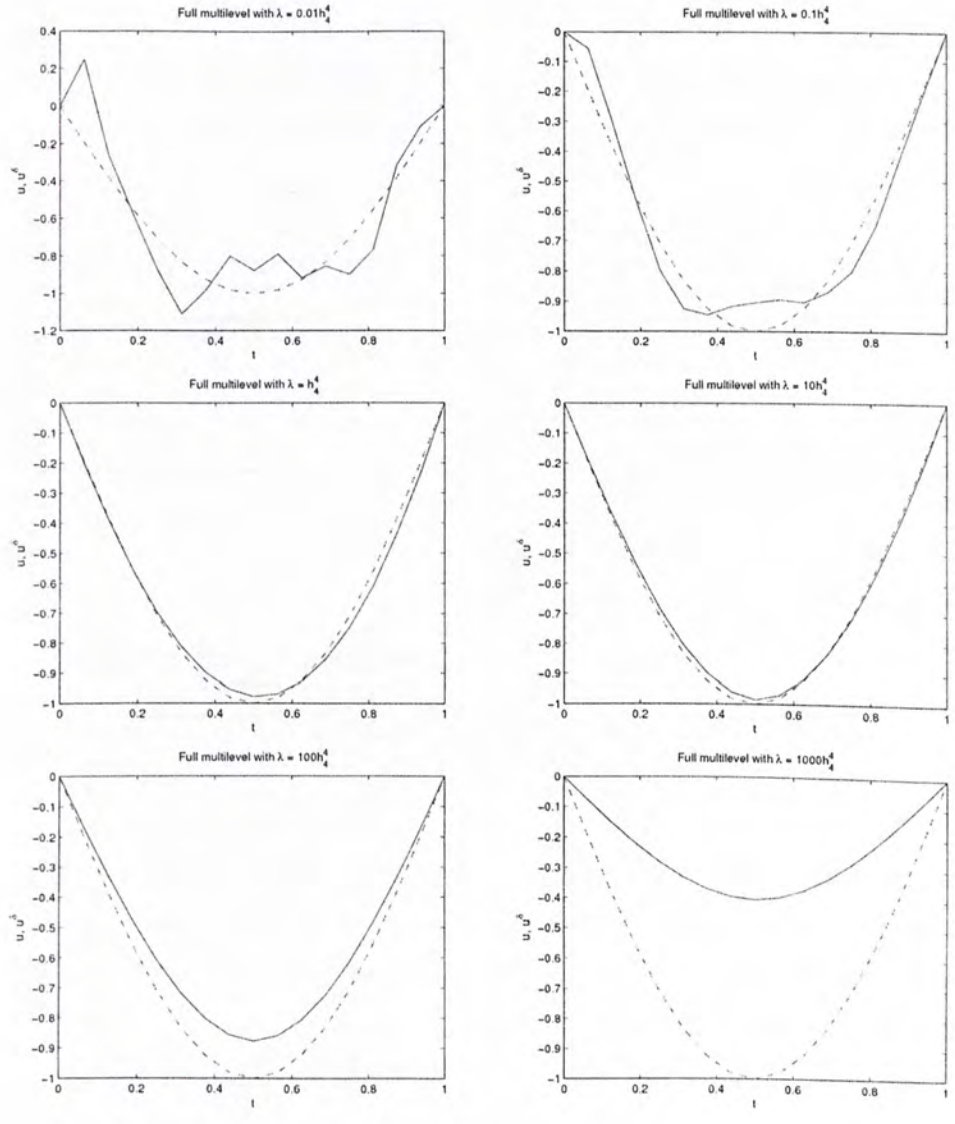

Figure 5.25: Full multilevel algorithm with different λ. The solid line represents the approximate solution with C_4 as the preconditioner, while the dashed line represents the exact solution. Here $\delta = 3\%$ and $\xi_j = (10^4\lambda + h_j^4)^{-1}$.

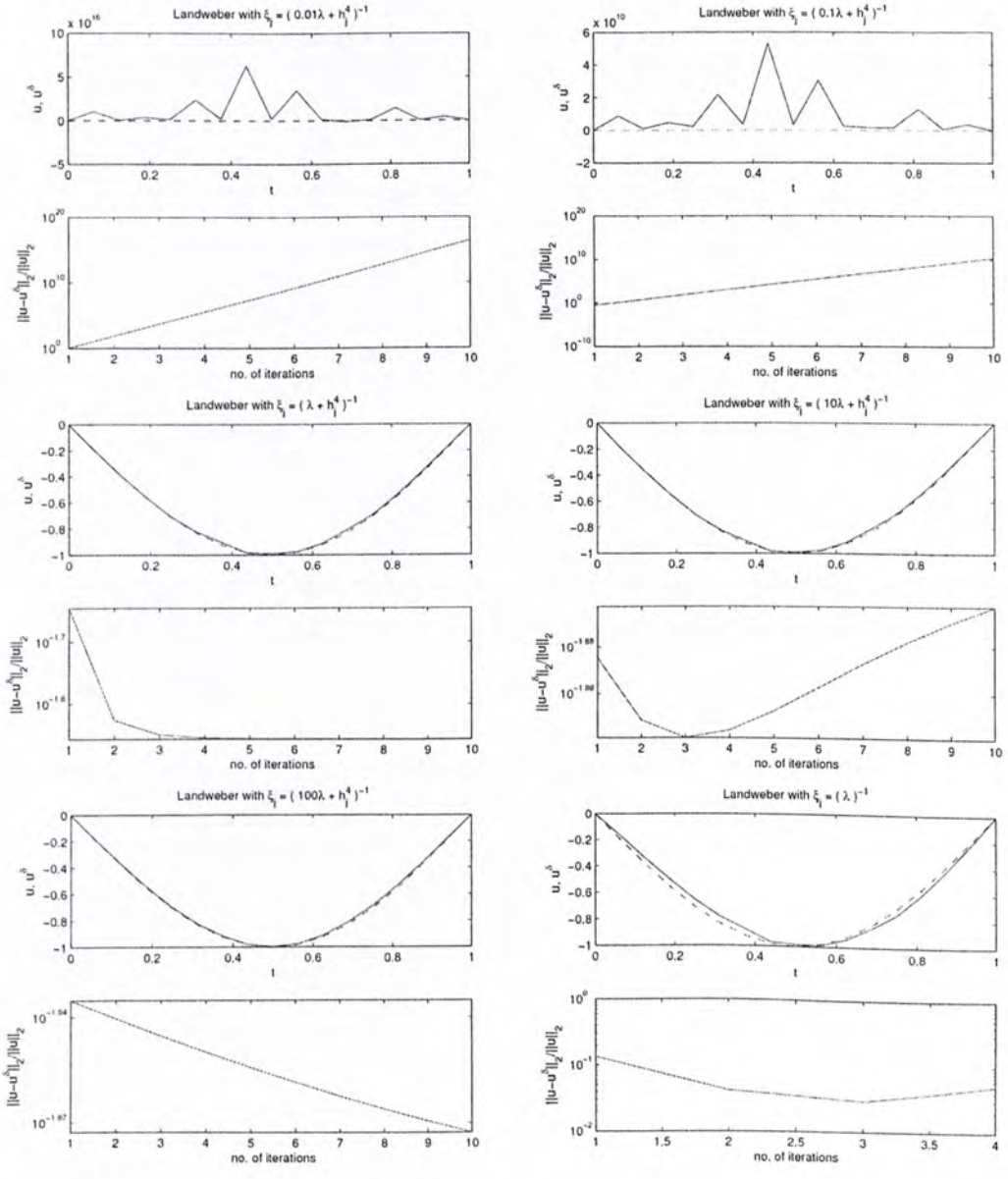

Figure 5.26: Nonzero regularization parameter - Landweber iteration with different ξ_j, and relative error. The solid line represents the approximate solution with C_4 as the preconditioner (10 iterations, except for the last $\xi_j = (\lambda)^{-1}$: 4 iterations), while the dashed line represents the exact solution. Here $\delta = 1\%$ and $\lambda = 10h_4^4$.

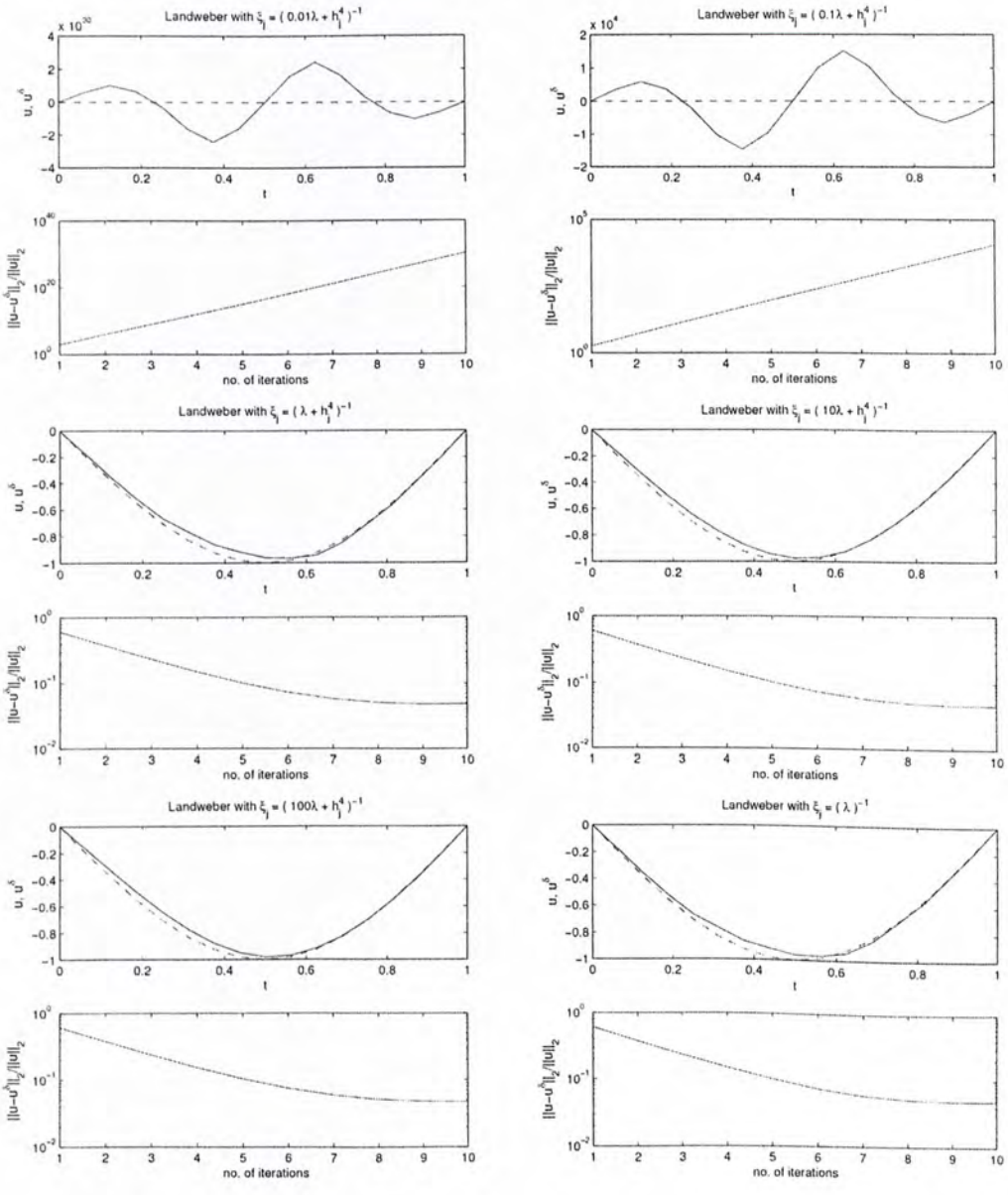

Figure 5.27: Zero regularization parameter - Landweber iteration with different ξ_j, and relative error. The solid line represents the approximate solution with C_4 as the preconditioner (10 iterations), while the dashed line represents the exact solution. Here $\delta = 10\%$ and $\lambda = 10^3 h_4^4$.

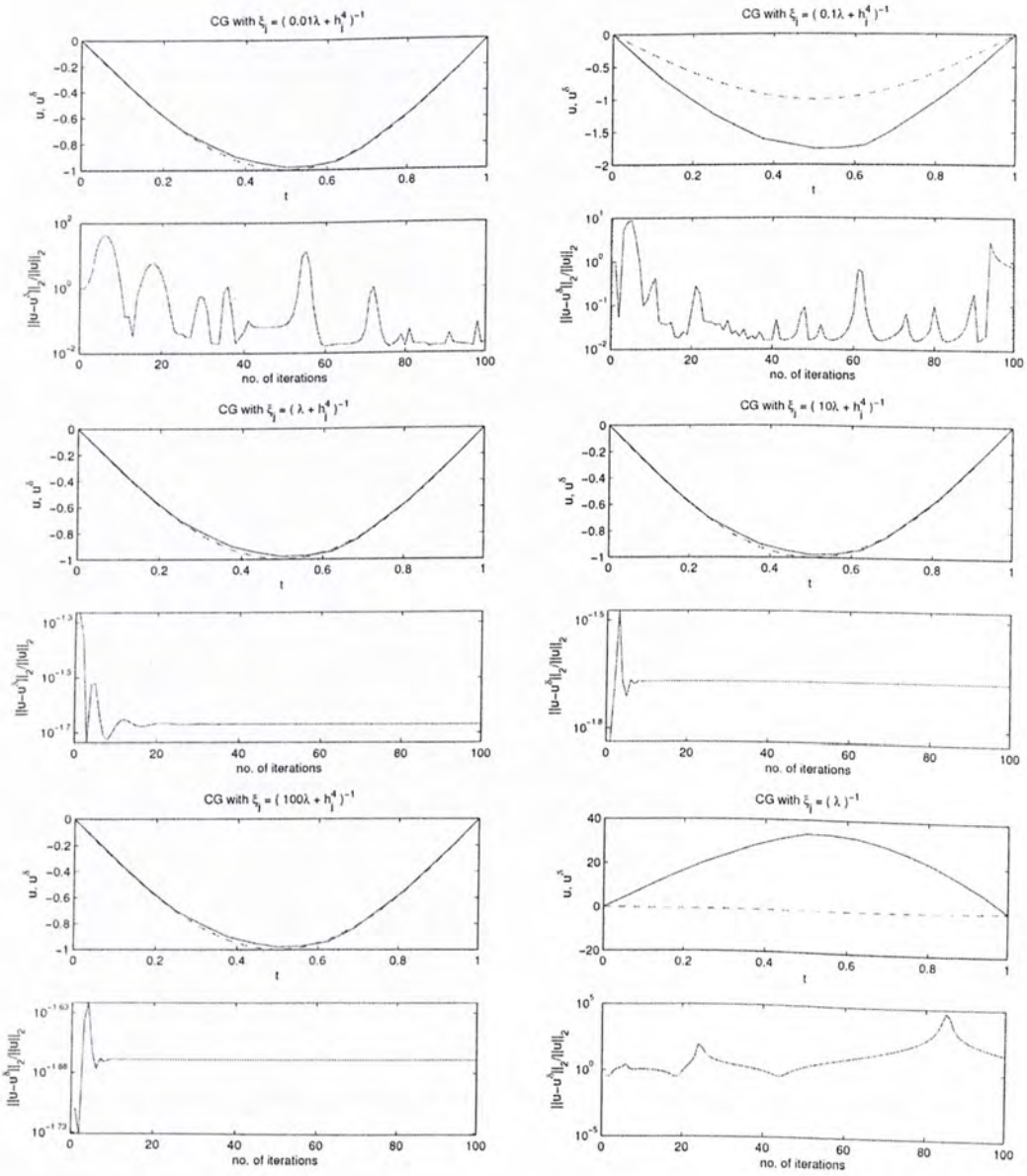

Figure 5.28: Nonzero regularization parameter - CG method with different ξ_j, and relative error. The solid line represents the approximate solution with C_4 as the preconditioner (100 iterations), while the dashed line represents the exact solution. Here $\delta = 1\%$ and $\lambda = 10h_4^4$.

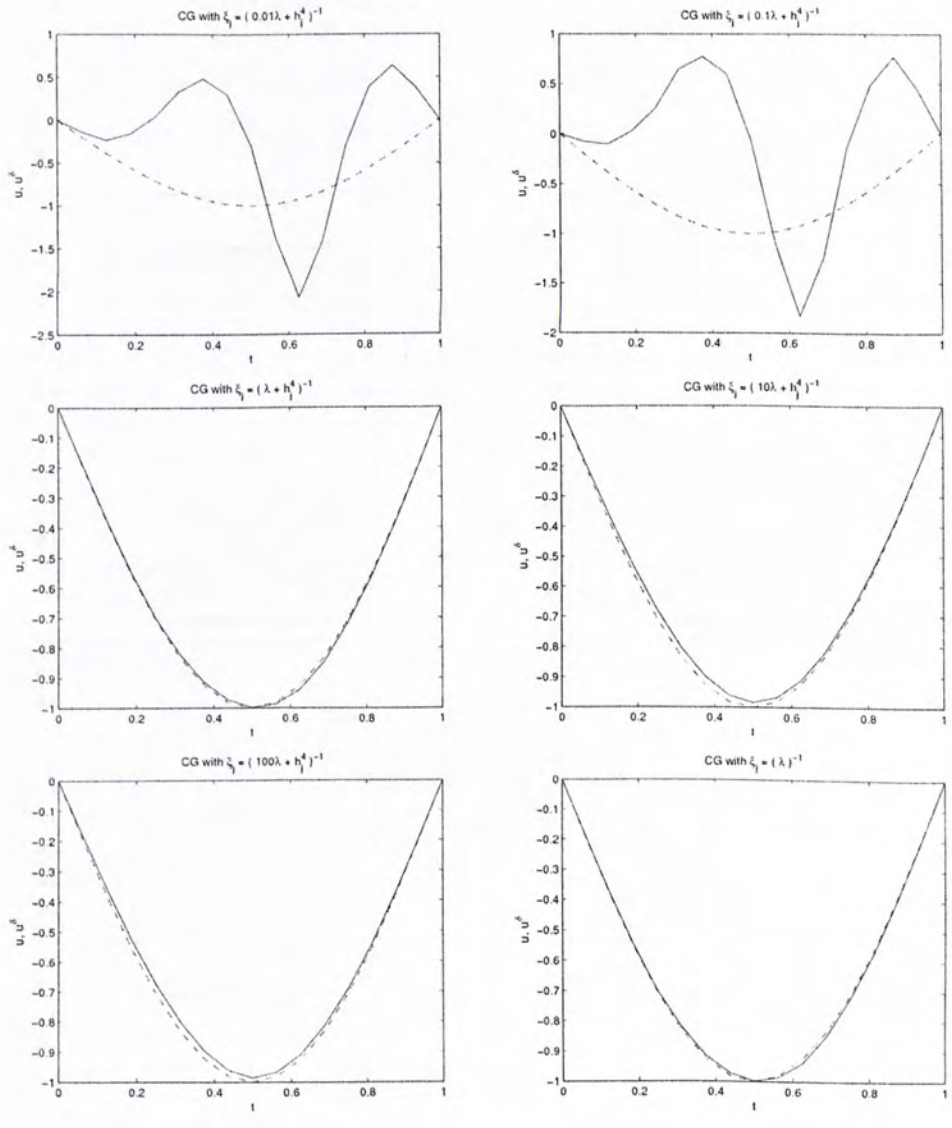

Figure 5.29: Zero regularization parameter - CG method with different ξ_j. The solid line represents the approximate solution with C_4 as the preconditioner (1 iteration), while the dashed line represents the exact solution. Here $\delta = 10\%$ and $\lambda = 10^3 h_4^4$.

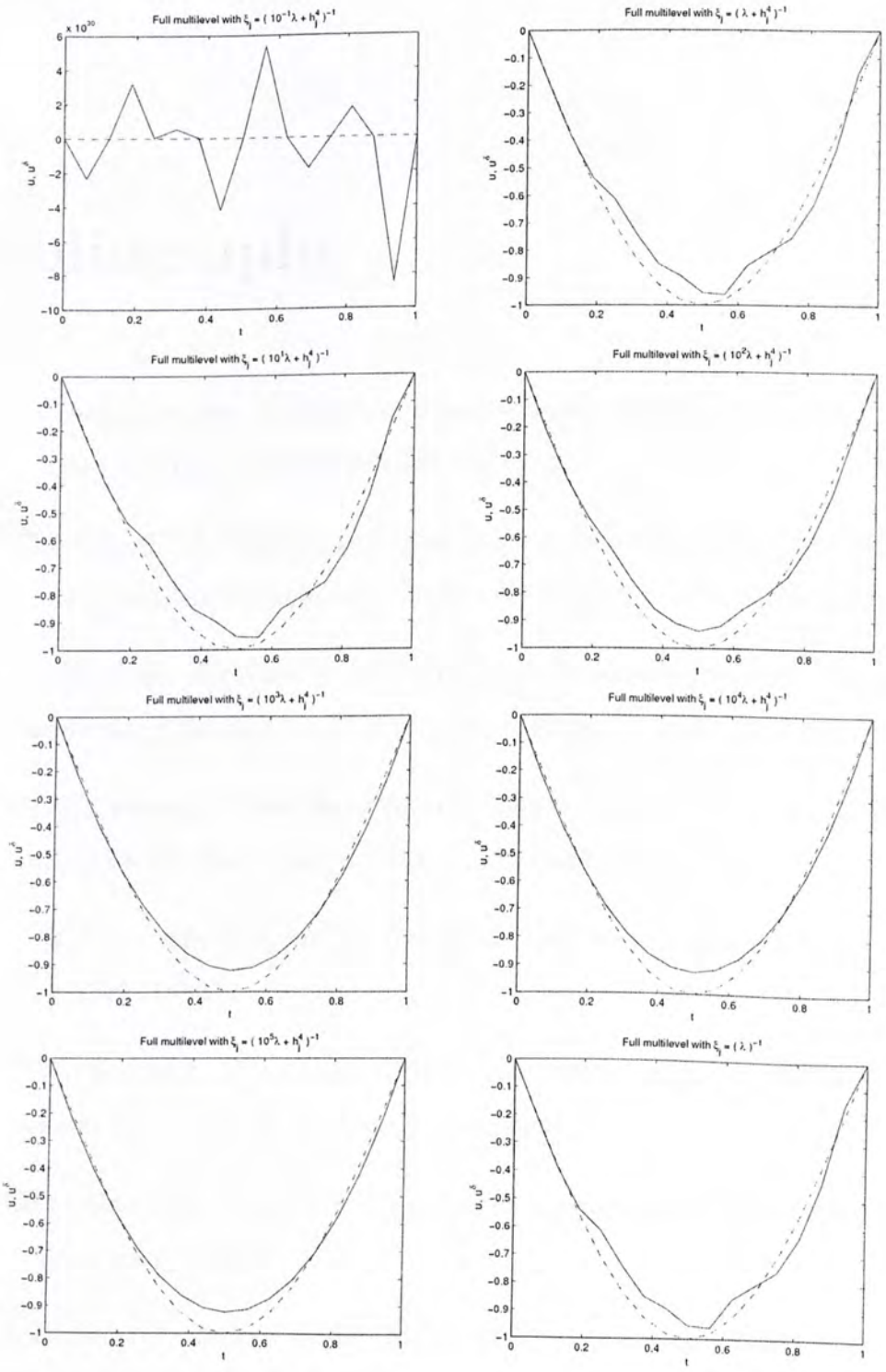

Figure 5.30: Full multilevel algorithm with different ξ_j. The solid line represents the approximate solution with C_4 as the preconditioner, while the dashed line represents the exact solution. Here $\delta = 10\%$ and $\lambda = 27h_4^4$.

Bibliography

[1] H. Bialy, *Iterative Behandlung linearer Funktionalgleichungen*, Arch. Rational Mech. Anal., 1 (1959/60), 166-176.

[2] H.W. Engl, M. Hanke, A. Neubauer, *Regularization of Inverse Problems*, Kluwer Academic Publishers, Dordrecht; Boston; London, 1996.

[3] V. Fridman, *Methods of successive approximations for Fredholm integral equations of the first kind*, Uspekhi Mat. Nauk, 11 (1956), 233-234.

[4] C.W. Groetsch, *Generalized Inverses of Linear Operators: Representation and Approximation*, Marcel Dekker, New York, 1977.

[5] C.W. Groetsch, *Elements of Applicable Functional Analysis*, Marcel Dekker, New York, 1980.

[6] C.W. Groetsch, *The theory of Tikhonov regularization for Fredholm equations of the first kind*, Pitman, Boston, 1984.

[7] C.W. Groetsch, *Inverse Problems in the Mathematical Sciences*, Vieweg, Braunschweig, 1993.

[8] J. Hadamard, *Lecture on the Cauchy Problem in Linear Partial Differential Equatiions*, Yale University Press, New Haven, 1923.

[9] P.C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems : numerical aspects of linear inversion*, SIAM, Philadelphia, 1998.

[10] B. Hofmann, *Regularization for Applied Inverse and Ill-posed Problems*, Teubner, Stuttgart, Germany, 1986.

[11] B. Kaltenbacher, *On the regularizing properties of a full multigrid method for ill-posed problems*, Inverse Problems, 17 (2001), 767-788.

[12] J.T. King, *A minimal error conjugate gradient method for ill-posed problems*, J. Optimization Theory Appl., 60 (1989), 297-304.

[13] J.T. King, *On the construction of preconditioners by subspace decomposition*, J. Comput. Appl. Math., 29 (1990), 195-205.

[14] J.T. King, *Multilevel algorithms for ill-posed problems*, Numer. Math., 61 (1992), 311-334.

[15] J.T. King, A. Neubauer, *A variant of Tikhonov regularization with a posteriori parameter choice*, Computing, 30 (1988), 91-109.

[16] A. Kirsch, *An Introduction to the Mathematical Theory of Inverse Problems*, Springer, New York, 1996.

[17] R. Kress, *Linear integral equations (2nd ed.)*, Springer, New York, 1999.

[18] Erwin Kreyszig, *Introductory functional analysis with applications*, Wiley, New York, 1978.

[19] L. Landweber, *An iteration formula for Fredholm integral equations of the first kind*, Amer. J. Math., 73 (1951), 615-624.

[20] A.K. Louis, *Inverse und schlecht gestellte Probleme*, Teubner-Verlag, Stuttgart, 1989.

[21] E. Schock, *On the asymptotic order of accuracy of Tikhonov regularization*, J. Optim. Theory Appl., 44 (1984), 95-104.

[22] M.H. Schultz, *Spline Analysis*, Prentice Hall, Englewood Cliffs, 1973.

[23] A.N. Tikhonov, *Regularization of incorrectly posed problems*, Soviet Math. Dokl., 4 (1963), 1624-1627.

[24] A.N. Tikhonov, *Solution of incorrectly formulated problems and the regularization method*, Soviet Math. Dokl., 4 (1963), 1035-1038.

[25] A.N. Tikhonov, V.Y. Arsenin, *Solutions of Ill-Posed Problems*, Wiley, New York, 1977.

[26] Curtis R. Vogel, *Computational Methods for Inverse Problems*, SIAM, Philadelphia, 2002.